



ETUDE CRISTALLOGRAPHIQUE DE LA PROTEINE IRP1 (IRON REGULATORY PROTEIN 1), UN REGULATEUR DE L'HOMÉOSTASIE DU FER.

Jérôme Dupuy

► To cite this version:

Jérôme Dupuy. ETUDE CRISTALLOGRAPHIQUE DE LA PROTEINE IRP1 (IRON REGULATORY PROTEIN 1), UN REGULATEUR DE L'HOMÉOSTASIE DU FER.. Biochimie [q-bio.BM]. Université Joseph-Fourier - Grenoble I, 2005. Français. NNT : . tel-00411037

HAL Id: tel-00411037

<https://theses.hal.science/tel-00411037>

Submitted on 25 Aug 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE

Présentée par

Jérôme DUPUY

*Pour obtenir le titre de docteur
de l'Université Joseph Fourier - Grenoble I
(arrêtés ministériels du 5 juillet 1984 et du 30 mars 1992)*

Spécialité : Cristallographie des macromolécules biologiques

Etude Cristallographique de la Protéine IRP1, un régulateur de l'homéostasie du fer.

*Soutenue le **31 mai 2005** devant le jury composé de*

Marc FONTECAVE

Christian DUMAS

Lukas KÜHN

Jean-Marc MOULIS

Anne VOLBEDA

Jean-Marc LHOSTE

Juan C. FONTECILLA-CAMPS

Président du Jury

Rapporteur

Rapporteur

Examineur

Examineur

Examineur

Directeur de Thèse

*Thèse préparée au
Laboratoire de Cristallographie et Cristallogenèse des Protéines
Institut de Biologie Structurale Jean-Pierre Ebel
CEA - CNRS - UJF*

Illustrations et Tableaux.....	5
Abréviations.....	10
Chapitre I : Introduction.....	12
1 <i>L'homéostasie du fer</i>	13
1.1 L'homéostasie cellulaire	14
1.1.1 L'incorporation du fer.....	14
1.1.2 Le stockage du fer.....	17
1.1.3 L'exportation du fer.....	18
1.1.4 La distribution intracellulaire du fer	19
1.2 L'homéostasie systémique.....	20
1.2.1 Les régulateurs systémiques	20
1.2.2 Accumulation, invasion et réactivité du fer	24
1.2.3 Le fer et les maladies neurodégénératives	26
1.2.3.1 La maladie de Parkinson	26
1.2.3.2 La maladie d'Alzheimer.....	29
1.2.3.3 Les autres maladies neurologiques.....	31
2 <i>Le système IRP/IRE</i>	32
2.1 IRP1 et IRP2.....	33
2.1.1 IRP1	34
2.1.1.1 Les protéines fer-soufre.....	36
2.1.1.2 Les homologues à IRP1	37
2.1.2 IRP2	38
2.2 Les cibles des IRP	39
2.3 Régulation des IRPs.....	43
2.3.1 Le stress oxydant	43
2.3.2 Le monoxyde d'azote (NO)	46
2.3.3 L'hypoxie.....	48
2.3.4 La croissance cellulaire.....	49
3 <i>Objectifs du travail de thèse</i>	50
Chapitre II : Matériel et méthodes	53
1 <i>La cristallogénèse</i>	53
1.1 Les principes.....	53
1.1.1 Les paramètres physico-chimiques	54
1.1.2 Les principaux agents précipitants.....	55
1.2 Les techniques	55
1.2.1 La diffusion en phase vapeur	55
1.2.2 L'ensemencement	56
1.3 La recherche des conditions de cristallisation	56
1.3.1 Le criblage	57
1.3.2 Les plans d'analyse factorielle.....	57
1.3.3 L'optimisation.....	58
1.4 La boîte à gants.....	59
1.4.1 Description.....	59
1.4.2 Manipulation dans la boîte à gants.....	60
1.5 Cryocristallographie et congélation des cristaux	60
1.5.1 Le propane liquide	61
1.5.2 Le rôle du cryoprotectant.....	61
2 <i>La cristallographie</i>	61
2.1 La diffraction des rayons X par les cristaux	62
2.1.1 Les directions de diffraction	62
2.1.2 Les facteurs de structure	63
2.2 Enregistrement des données	64
2.2.1 La ligne de lumière FIP – BM30a.....	65
2.2.2 Les lignes de lumière ID14.....	66

2.3 Le traitement des données	66
2.3.1 Les principes	66
2.3.2 Les programmes utilisés	67
2.3.3 La validation des données	68
2.4 Caractéristiques des cristaux de protéines	69
2.4.1 Régularité du réseau cristallin ou mosaïcité	69
2.4.2 Désordre des atomes du cristal	69
2.5 Analyse du contenu de l'unité asymétrique	70
2.5.1 Estimation du nombre de molécules de l'unité asymétrique (Z_a)	71
2.5.2 Identification des symétries non cristallographiques	72
2.6 Méthodes de phasage	73
2.6.1 Le problème des phases	73
2.6.2 Remplacement moléculaire	73
2.6.2.1 Fonction de rotation	74
2.6.2.2 Fonction de translation	74
2.6.2.3 Les conditions	75
2.6.3 Utilisation d'atomes lourds	75
2.6.3.1 Méthode MIR	75
2.6.3.1.1 Atomes lourds utilisés et modes de fixation	78
2.6.3.2 Méthode MAD	78
2.6.3.2.1 La diffusion anormale	79
2.6.3.2.1.1 Principe	79
2.6.3.2.1.2 Obtention du facteur de diffusion anormale	80
2.6.3.2.2 Aspects théoriques de la méthode MAD	81
2.6.3.2.4 Protocole de mesures	83
2.6.3.2.4.1 Rapport anomal et « dispersif »	83
2.6.3.2.4.2 Acquisition de données MAD	83
2.6.3.2.5 Choix des diffuseurs anormaux	85
2.6.3.3 Cas particulier du SAD	86
2.6.4 Méthodes directes	87
2.7 Modification de densité	88
2.7.1 Principe de l'aplatissement de solvant	88
2.7.2 Symétries non cristallographiques – moyennation	89
2.7.2.1 Symétries non cristallographiques	89
2.7.2.2 Moyennation	89
2.7.3 Procédure de modification de densité électronique	89
2.8 Affinement et construction de la structure	90
2.8.1 L'affinement par minimisation	91
2.8.1.1 Le terme à minimiser	91
2.8.1.2 Les moindres carrés	92
2.8.1.3 Le maximum de vraisemblance	92
2.8.1.4 Le terme d'énergie	93
2.8.2 La dynamique moléculaire	93
2.8.3 Les programmes utilisés	94
2.8.3.1 REFMAC	94
2.8.3.2 CNS	94
2.8.4 Cartes de densité électronique	95
2.8.5 Critères d'évaluation du modèle	96
Chapitre III : Résultats	99
1 Les premières données	99
1.1 Cristallisation	99
1.2 Cristaux natifs	100
1.2.1 Collecte des données et statistiques	100
1.2.2 Nombre de molécules par unité asymétrique	101
1.3 Enregistrement de jeux MAD	102
1.4 Détermination de la position des agrégats dans l'unité asymétrique	104
2 Résolution du problème de la phase	105
2.1 Cristallisation	105

2.2 Dérivés d'atomes lourds	107
2.3 Phasage par la méthode MAD	108
2.3.1 Mesure de fluorescence et choix des longueurs d'onde	108
2.3.2 Intégration et mise à l'échelle – statistiques	110
2.3.3 Nombre de molécules par unité asymétrique	111
2.3.4 Détermination de la position des agrégats et des atomes lourds	111
2.3.5 Phasage et carte de densité électronique	113
2.3.6 Construction du premier modèle	116
3 Résolution de la structure	118
3.1 Amélioration des cartes de densité électronique	118
3.2 Nouveau modèle moléculaire	120
3.3 Remplacement moléculaire	120
3.3.1 Forme à haute résolution monoclinique P2 ₁	120
3.3.2 Forme orthorhombique C222 ₁	123
3.4 Affinement	124
Chapitre IV : Analyse de la structure d'IRP1	127
1 <i>IRP1</i> versus l'aconitase mitochondriale (<i>mAcn</i>)	127
1.1 Description générale	127
1.2 Analyse par domaine	129
1.3 L'activité aconitase d'IRP1	132
1.3.1 Site actif	132
1.3.2 Le mécanisme catalytique	136
1.3.3 Tunnel d'accessibilité	140
1.3.4 Le site zinc	141
1.3.5 Sérines phosphorylables	142
2 Modélisation de la fixation du motif IRE	145
3 <i>IRP1</i> versus <i>IRP2</i>	159
Chapitre V : Conclusion et perspectives	165
Bibliographie	168

Illustrations et Tableaux

Illustrations

Chapitre I : Introduction

Figure I.1 : Le cycle de la transferrine.	15
Figure I.2 : Métabolisme cellulaire du fer.	17
Figure I.3 : Homéostasie systémique du fer.	22
Figure I.4 : Rôle de l'hépcidine, de la protéine HFE et d'autres molécules impliquées dans la régulation de l'homéostasie systémique du fer.	23
Figure I.5 : Localisation des différentes zones du cerveau présentant des anomalies quant à leur niveau de concentration en fer.	25
Figure I.6 : Schéma présentant les mécanismes à la base de la maladie de Parkinson.	29
Figure I.7 : Structure de l'aconitase mitochondriale de bœuf (code pdb 1C96).	33
Figure I.8 : Régulation d'IRP1.	35
Figure I.9 : Régulation d'IRP2.	39
Figure I.10 : Motif IRE dans les ARNm.	39
Figure I.11 : Régulation de la traduction par le système IRP/IRE.	40
Figure I.12 : Régulation de la traduction par le système IRP/IRE.	42
Figure I.13 : Système de régulation IRP/IRE.	42
Figure I.14 : Régulation redox d'IRP1 et de l'aconitase mitochondriale.	45
Figure I.15 : Modifications post-transcriptionnelles dans le métabolisme du fer des macrophages lors des phénomènes inflammatoires.	47

Chapitre II : Matériel et méthodes

Figure II.1 : Différents chemins expérimentaux sur le diagramme de phase.	54
Figure II.2 : Technique de la goutte suspendue.	56
Figure II.3 : Limitation de la diffusion de vapeur grâce à l'emploi d'une barrière minérale constituée d'un mélange d'huiles plus ou moins perméable à l'eau.	59
Figure II.4 : Vue de la boîte à gants.	60
Figure II.5 : Schéma décrivant le phénomène de la diffraction selon les lois de Laue (A) et de Bragg (B).	63
Figure II.6 : Système CATS (Cryogenic Automated Transfert System).	65

Figure II.7 : Cristal généré à partir de l'unité asymétrique.	71
Figure II.8 : Représentation du facteur de structure de l'ensemble protéine-atome lourd.	76
Figure II.9 : Construction de Harker.	77
Figure II.10 : Facteurs anomaux près du seuil d'absorption K du sélénium ($\lambda=0,98$ Å) dans un cristal de thiorédoxine sélénée de <i>E. coli</i> .	79
Figure II.11 : Variation de la partie imaginaire f'' du facteur de diffusion atomique pour différents complexes sélénium, au seuil d'absorption K.	81
Figure II.12 : Représentation vectorielle montrant le non respect de la loi de Friedel en présence d'un diffuseur anomal.	82
Figure II.13 : Variation des facteurs f' et f'' .	84
Figure II.14 : Schéma de la procédure de modification de densité électronique.	90
 Chapitre III : Résultats	
Figure III.1 : Evolution de la morphologie des cristaux.	100
Figure III.2 : Spectre de fluorescence au seuil d'excitation du fer.	102
Figure III.3 : Spectre de fluorescence au seuil K du fer mesuré sur le cristal afin de tenir compte de l'influence des ligands.	103
Figure III.4 : Carte de Patterson du cristal 40, section de Harker $v=1/2$.	104
Figure III.5 : Evolution au cours du temps de la morphologie des cristaux.	106
Figure III.6 : Stabilisation des cristaux avec différents solvants.	106
Figure III.7 : Rôle de la dilution sériée dans l'ensemencement sur le contrôle du nombre de noyau de nucléation.	107
Figure III.8 : Limitation de la diffusion de vapeur.	107
Figure III.9 : Spectre de fluorescence au seuil d'excitation de l'or.	109
Figure III.10 : Seuil d'absorption des différents diffuseurs anomaux.	109
Figure III.11 : Carte de Patterson de la différence anormale de l'agrégat.	111
Figure III.12 : Fichier de sortie de SHELXD.	112
Figure III.13 : Projection suivant l'axe b de la densité électronique.	114
Figure III.14 : Evolution de la figure de mérite en fonction de la résolution.	115
Figure III.15 : Extrait du fichier de sortie du programme RESOLVE.	115
Figure III.16 : Obtention du premier modèle.	117
Figure III.17 : Structure incomplète d'IRP1.	117

Figure III.18 : Récapitulatif des figures de mérite.	118
Figure III.19 : Projection suivant l'axe b de la densité électronique.	119
Figure III.20 : Amélioration de la carte de densité électronique.	119
Figure III.21 : Amélioration du modèle.	120
Figure III.22 : Structure incomplète d'IRP1 dans le cristal 16.	121
Figure III.23 : Suivi de l'évolution de la carte de densité électronique.	122
Figure III.24 : Carte de densité électronique d'une région difficilement interprétable.	123
Figure III.25 : Structure finale d'IRP1.	123
Figure III.26 : Diagramme de Ramachandran, correspondant à la molécule IRP1 de l'unité asymétrique, calculé avec le programme PROCHECK	125

Chapitre IV : Analyse de la structure d'IRP1

Figure IV.1 : Réaction d'isomérisation du citrate en isocitrate avec son intermédiaire réactionnel, le <i>cis</i> -aconitate.	127
Figure IV.2 : Structure de la protéine IRP1.	128
Figure IV.3 : Représentation de l'organisation des éléments de structures secondaires dans les domaines 1, 2 et 3.	129
Figure IV.4 : Représentation de l'organisation des éléments de structures secondaires dans le domaine 4.	130
Figure IV.5 : Organisation des feuillettes de brins β parallèles.	130
Figure IV.6 : Conservation des éléments de structure de mAcn chez IRP1.	132
Figure IV.7 : Alignement de séquences entre la protéine IRP1 humaine, IRP2 et l'aconitase mitochondriale.	133
Figure IV.8 : Représentation stéréoscopique du site actif d'IRP1.	134
Figure IV.9 : Représentation stéréoscopique des acides aminés impliqués dans des interactions avec l'agrégat fer soufre.	135
Figure IV.10 : Représentation stéréoscopique des acides aminés impliqués dans des contacts avec le substrat.	136
Figure IV.11 : Modèle de mécanisme réactionnel catalysé par IRP1 et l'aconitase.	137
Figure IV.12 : Représentation schématique de la réaction transformant l'isocitrate en <i>cis</i> -aconitate.	139
Figure IV.13 : Représentation des deux chemins d'accès à l'agrégat depuis la surface de la protéine.	140

Figure IV.14 : Vue des deux molécules symétriques mettant en commun les acides aminés nécessaires à la coordination des atomes de zinc.	141
Figure IV.15 : Sites d'interactions entre les molécules symétriques.	142
Figure IV.16 : Localisation des deux sérines phosphorylables sur la protéine IRP1.	142
Figure IV.17 : Environnement de la sérine 138.	143
Figure IV.18 : Représentation des acides aminés indispensables au site actif et qui sont aussi à proximité de la sérine 138.	143
Figure IV.19 : Localisation de l'arginine 713 par rapport à la sérine 711.	144
Figure IV.20 : Structure secondaire du motif IRE de l'ARN messager de la ferritine.	146
Figure IV.21 : Représentation de la localisation des acides aminés impliqués dans les interactions avec l'IRE.	148
Figure IV.22 : Représentation des zones de contacts entre les domaines 1, 2, 3 et le domaine 4 sans le <i>linker</i> .	149
Figure IV.23 : Localisation des acides aminés supposés comme étant impliqués dans la fixation des motifs IRE.	151
Figure IV.24 : Superposition du motif IRE et du domaine 4.	152
Figure IV.25 : Représentation des deux orientations possibles pour le positionnement de l'IRE à la surface des domaines 1, 2 et 3.	152
Figure IV.26 : Positionnement du motif IRE forme « tête haute » sur les domaines 1, 2 et 3.	153
Figure IV.27 : Modèle final avec un IRE en position tête haute.	154
Figure IV.28 : Localisation du patch hydrophobe constitué par le domaine 1 et le domaine 4.	155
Figure IV.29 : Localisation de la zone hydrophobe constituée par les domaines 1 et 4 après la modélisation de la fixation du motif IRE.	156
Figure IV.30 : Positionnement du motif IRE sous un forme tête basse.	156
Figure IV.31 : Superposition de l'hélice $\alpha 30$ avant et après modélisation.	157
Figure IV.32 : Modélisation de la fixation de l'IRE sur IRP1.	157
Figure IV.33 : Modélisation de la nouvelle position du <i>linker</i> .	158
Figure IV.34 : Alignement de séquences entre la protéine IRP1 humaine et IRP2.	160
Figure IV.35 : Localisation de la sérine 441 et de l'isoleucine 176 à proximité de l'agrégat fer soufre.	161

Figure IV.36 : Localisation de la zone d'insertion des 73 résidus supplémentaires de IRP2 sur la structure d'IRP1.	162
--	-----

Tableaux

Chapitre II : Matériel et méthodes

Tableau II.1 : Systèmes cristallins et groupes d'espace.	70
--	----

Chapitre III : Résultats

Tableau III.1 : Paramètres de maille des cristaux natifs.	100
Tableau III.2 : Statistiques des données natives indexées et intégrées par XDS.	101
Tableau III.3 : Volume de Matthews correspondant aux deux jeux natifs.	101
Tableau III.4 : Statistiques des données MAD après leur mise à l'échelle.	103
Tableau III.5 : Paramètres de maille des différentes indexations possibles pour les cristaux 40 et 16.	104
Tableau III.6 : Statistiques des jeux natifs après indexation dans les nouveaux groupes d'espace.	105
Tableau III.7 : Sels d'atomes lourds testés et concentrations utilisées.	108
Tableau III.8 : Paramètres des différentes longueurs d'onde utilisées lors des collectes et de leurs facteurs f^2 et f^2 mesurés.	110
Tableau III.9 : Paramètres de maille des deux cristaux 1 et 2 juste après l'indexation.	110
Tableau III.10 : Statistiques des jeux MAD et SAD collectés sur les cristaux 1 et 2.	110
Tableau III.11 : Statistiques du phasage en sortie de SHARP.	113
Tableau III.12 : Statistiques en fin d'affinement avec le programme REFMAC.	124

Chapitre IV : Analyse de la structure d'IRP1

Tableau IV.1 : Tableau récapitulatif des écarts types entre les domaines d'IRP1 et ceux de l'aconitase mitochondriale.	131
Tableau IV.2 : Accessibilité des acides aminés aromatiques impliqués dans des contacts entre le domaine 4 et les domaines 1, 2 et 3 sur une distance inférieure à 5 Å.	150

Abréviations

$\langle I/\sigma \rangle$	signal sur bruit moyen
$^{\circ}\text{C}$ / K	degré celcius / Kelvin
Å	Angström
a,b,c	vecteurs unité de la maille cristalline
B	facteur d'agitation thermique
Cp	Céruloplasmine
Da / kDa	Dalton / kilodalton
eV	Electron volt
ESRF	European Synchrotron Radiation facility
FIP	<i>French beamline for Investigation of Proteins</i>
Ft	Ferritine
F / F	Facteur de structure / module de F
h, k, l	Indices de Miller
IRE	<i>Iron Responsive Element</i>
IRP	<i>Iron Regulatory Protein</i>
l / ml / μl	litre / millilitre / microlitre
mAcn	Aconitase mitochondriale de boeuf
MAD	<i>Multiwavelength Anomalous Diffraction</i>
MIR	<i>Multiple Isomorphous Replacement</i>
M / mM	molaire / millimolaire
mg/ml	milligramme par millilitre
mg / μg	milligramme / microgramme
mm / μm	millimètre / micromètre
PDB	Protein Data Bank
PEG	Polyéthylène Glycol
r.m.s.	Root Mean Scare (déviation moyenne)
$R/R_{\text{work}}/R_{\text{free}}/R_{\text{cryst}}$	coefficients d'accord entre les facteurs de structure observés et calculés
$R_{\text{sym}}/R_{\text{meas}}$	coefficient d'accord entre réflexions symétriques / pondéré
SAD	<i>Single Anomalous Diffraction</i>
SIRAS	<i>Single Isomorphous Replacement with Anomalous Diffraction</i>
Tf	Transferrine
TfR	Récepteur à la transferrine
Tris	Tris(hydroxyméthyl)aminomethane
V	volume
V_{m}	coefficient de Matthews
x, y, z	coordonnées atomiques
Z	nombre de positions équivalentes dans la maille

Introduction

Les devises Shadok



QUAND ON NE SAIT PAS OÙ L'ON VA,
IL FAUT Y ALLER !!...
... ET LE PLUS VITE POSSIBLE.

Chapitre I : Introduction

Ce travail s'intègre dans le cadre de l'étude du système de régulation IRP/IRE, pour *Iron Regulatory Proteins* et *Iron Responsive Element*, qui contrôle l'homéostasie du fer chez l'homme. Après avoir énoncé les grands principes de cette homéostasie, nous aborderons, par un descriptif non exhaustif, les problèmes pathologiques majeurs engendrés par une dérégulation du contrôle de la concentration en fer. Nous décrirons ensuite, et par le détail, le rôle que joue le système IRP/IRE sur cette régulation et les différents mécanismes mis en place pour infléchir son fonctionnement. Enfin nous conclurons par les choix que nous avons faits pour aborder ce sujet et les moyens que nous avons mis en place pour répondre à un certain nombre de questions qui se posaient alors.

L'homéostasie des métaux est gouvernée par deux conséquences intrinsèques à l'évolution. La première concerne les réactions d'oxydoréductions qui sont des processus indispensables à la vie. Les métaux de transition sont des éléments essentiels au fonctionnement des nombreuses protéines impliquées dans le contrôle de ce type de réactions. Ceci entraîne donc une compétition entre les cellules pour le contrôle de ces ressources en ions métalliques. La deuxième conséquence est liée au fait que plusieurs voies métaboliques induisent la synthèse d'agents potentiellement toxiques lorsque les ions métalliques sont présents en quantités anormales. Cela a pour effet d'induire des dommages aussi bien vis-à-vis de la fonction même de la voie métabolique que sur les protéines et les acides nucléiques présents dans leur environnement direct.

Les différents ions métalliques peuvent être regroupés en deux classes, ceux possédant une activité d'oxydoréduction dans un contexte biologique comme ce peut être le cas du Fe^{2+} , du Cu^{2+} , du Co^{2+} , du Ni^{2+} et du Mn^{2+} et ceux qui en sont dépourvus comme pour le Ca^{2+} , le Mg^{2+} et le Zn^{2+} . Le calcium, le magnésium et le zinc sont souvent associés à des facteurs de transcription ou à des enzymes impliqués dans le métabolisme de l'ADN, à des fins de reconnaissance ou de cofacteurs, car la présence d'un ion métallique ayant une activité redox pourrait être à l'origine de réactions radicalaires qui auraient pour finalité la dégradation des acides nucléiques. Tous ces processus requièrent donc une quantité définie d'ions métalliques à des moments bien précis du développement cellulaire et pour des fonctions parfaitement contrôlées et régulées.

Bien que le fer soit un élément abondant sur notre planète, sa disponibilité à la surface de la terre est limitée (10^{-18} M). Cette limitation globale en fer est la résultante de notre

atmosphère riche en oxygène, qui contraint le fer sous sa forme Fe^{3+} . Le fer ferrique (Fe^{3+}) est insoluble à pH neutre et basique, ce qui le rend indisponible pour les organismes vivants. Pour lutter contre cela les eucaryotes, mais aussi les procaryotes, ont développé des systèmes extrêmement efficaces afin de pouvoir récupérer le peu de fer disponible dans leur environnement et ce jusqu'à des concentrations extrêmement faibles.

1 L'homéostasie du fer

Environ 20 mg de fer sont nécessaires pour renouveler les quelques 200 milliards de globules rouges produits quotidiennement, ce qui représente près de 80 % de la demande en fer chez un adulte normalement constitué. Il n'est pas surprenant, alors, qu'un déficit en fer conduise à des anémies, ce qui place cette maladie au cœur des problèmes de santé publique car elle touche pratiquement un quart de la population mondiale. Ainsi des cellules carencées en fer arrêtent de proliférer et peuvent finir par mourir si cet état perdure.

Le fer a une structure électronique qui lui permet d'avoir préférentiellement deux états d'oxydation différents, Fe^{2+} et Fe^{3+} , le rendant de ce fait propice à son utilisation en temps que cofacteur dans de nombreuses réactions d'oxydoréduction. Cependant, cette même propriété chimique explique pourquoi un excès de fer libre peut être toxique. Dans le cytoplasme, une partie significative du fer est sous forme réduite et peut donc participer à un type de réaction chimique qui est nommée communément, réaction de Fenton :



le fer ferreux réagit avec le peroxyde d'hydrogène (H_2O_2), ou des lipides peroxydés, pour générer du fer ferrique, des ions hydroxydes (OH^-), et des radicaux hydroxyles ($\text{OH}\cdot$) qui sont eux très réactifs (ou des radicaux lipidiques comme $\text{LO}\cdot$ et $\text{LOO}\cdot$). L'ensemble de ces radicaux représente alors un risque pour l'intégrité des lipides membranaires, des protéines et des acides nucléiques. Ainsi une surcharge cellulaire en fer, aussi bien qu'un déficit, peut conduire la cellule à une mort certaine, c'est pourquoi la concentration en fer doit être soigneusement contrôlée et limitée.

Beaucoup de pathologies résultent d'une accumulation de fer dans les tissus mais une régulation systémique, c'est-à-dire au niveau de l'organisme, n'est pas une condition suffisante pour éviter ce genre de problème. En effet, des dérégulations locales dans l'homéostasie du fer contribuent aussi à la genèse de désordres neurodégénératifs comme ce peut-être le cas dans la maladie de Parkinson. Il existe donc différents mécanismes, que ce soit au niveau d'une simple cellule, d'un tissu ou encore d'un organisme entier, pour

permettre d'équilibrer cette balance qui risque à tout moment de basculer du déficit à l'excès de fer, les deux ayant une finalité plutôt néfaste.

1.1 L'homéostasie cellulaire

L'homéostasie cellulaire du fer nécessite un contrôle fin de l'incorporation, du stockage, de l'exportation et de la gestion intracellulaire des ressources en fer. Bien que les organismes unicellulaires doivent faire face à ce même type de challenge, les vertébrés ont développé différents mécanismes hautement spécialisés pour répondre à cette demande.

1.1.1 L'incorporation du fer

Chez un individu en parfaite santé, le fer extracellulaire circule dans le plasma de façon complexée à la transferrine (Tf). Il s'agit d'une protéine abondante qui a une extraordinaire affinité pour le fer. La cellule a mis en place différents mécanismes pour internaliser les complexes de Tf-fer (Figure I.1). Un des modèles les mieux étudiés est celui du récepteur 1 de la transferrine (TfR1) qui fixe la transferrine di-ferrique (Cheng et al., 2004). Les complexes Tf/TfR1 se forment au niveau des puits recouverts de la membrane plasmique et sont rapidement absorbés dans la cellule par endocytose. En rejoignant l'endosome, qui est acide (pH entre 5 et 6), le fer ferrique se dissocie de la transferrine. L'acidité abaisse l'affinité de la transferrine pour le Fe^{3+} d'environ un million de fois. L'acidification provoque la libération du fer par protonation des liaisons qui contribuent à sa fixation. Une ferriréductase, actuellement toujours non identifiée, réduit le fer ferrique en fer ferreux permettant ainsi au transporteur de métaux divalent (DMT-1, connu aussi sous les noms de DCT1 et Nramp2) de transférer le Fe^{2+} à travers la membrane de l'endosome vers le cytoplasme (Fleming et al., 1998). L'apoTf et le TfR1 sont ensuite recyclés à la surface de la cellule, où chacun peut être réutilisé pour un nouveau cycle. Du fait que le TfR1 est constitutivement exprimé, l'incorporation du fer via la transferrine se produit dans beaucoup de types cellulaires (Figure I.1). Cependant, cette voie est prépondérante dans le développement de la lignée érythroblastique, aboutissant à la synthèse des érythrocytes (ou hématies), car il y a une énorme demande en fer du fait même de la synthèse de l'hémoglobine.

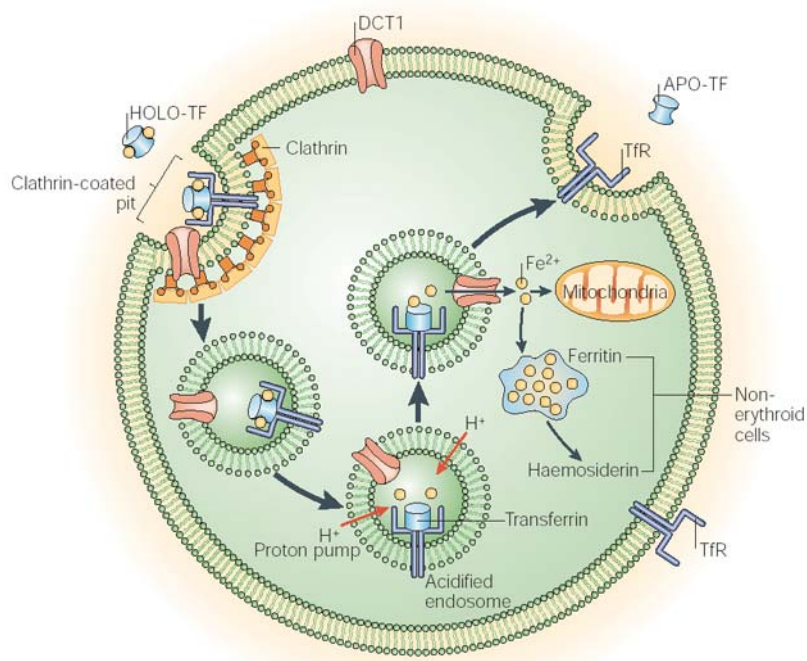


Figure I.1 : Le cycle de la transferrine. L'holotransferrine (HOLO-TF) se lie au récepteur de la transferrine (TfR1) à la surface de la cellule. Ces complexes, localisés dans les zones riches en puits recouverts, sont internalisés par endocytose. L'endosome acidifie le milieu, grâce à la présence de pompes à proton au niveau de sa membrane, permettant ainsi le relargage du fer qui est alors transporté vers le milieu cellulaire grâce à DMT-1 (ou DCT1). L'apotransferrine (APO-TF) retourne vers la membrane, toujours complexée au TfR1, où, à pH neutre, ils se dissocient pour participer de nouveau au transport du fer. Le fer peut être ensuite adressé à la mitochondrie ou alors il est stocké grâce à la ferritine et à l'hémossidérine. D'après Zecca et al. (2004).

Une protéine homologue, le récepteur à la transferrine 2 (TfR2) (Kawabata et al., 1999), est retrouvée, de façon spécifique, sur les hépatocytes et les cellules érythroïdes, suggérant un rôle plus spécialisé. La découverte de mutations sur le gène humain codant pour le TfR2 a abouti à la caractérisation d'une maladie connue sous le nom d'hémochromatose (Camaschella et al., 2000). Il s'agit d'une surcharge en fer au niveau de tout l'organisme. Cette découverte a permis d'établir l'importance du TfR2 dans l'homéostasie du fer. Le TfR2 lie la transferrine avec, approximativement, trente fois moins d'affinité que le TfR1, mais contrairement au TfR1, son expression n'est pas contrôlée par le système de régulation IRE/IRP (voir § I.2). Dans des cellules hautement spécialisées, comme c'est le cas pour les cellules épithéliales du rein, il existe un troisième mécanisme d'incorporation du fer fixé à de la transferrine, qui n'utilise plus un récepteur à la transferrine mais profite de la présence d'un récepteur cubiline/mégaline (Kozyraki et al., 2001). Bien que l'incorporation du fer via le système Tf soit prédominant dans des conditions normales, une surcharge en fer conduit à la saturation de ce transporteur et ainsi à la circulation de fer libre. Certains indices laissent à penser qu'il existe plusieurs autres voies de transport pour le fer non lié à la transferrine

Ainsi l'absorption du fer par les cellules intestinales se fait à travers l'action du DMT-1, le même transporteur de métaux qui régule aussi l'exportation du fer de l'endosome après que

celui-ci ait été internalisé via le cycle Tf/TfR1 (Fleming et al., 1997 ; Gunshin et al., 1997). DMT-1 est un symporteur proton/métal divalent qui n'est pas spécifique que du fer, ce qui lui permet d'échanger d'autres métaux tels que du Mn^{2+} , Cu^{2+} , Zn^{2+} Des études réalisées en culture cellulaire montrent qu'il y a au moins un autre système direct d'incorporation du fer distinct de celui de DMT-1 mais il reste encore mal caractérisé (Sturrock et al., 1990). Les canaux calciques peuvent, eux aussi sous certaines conditions, réguler l'importation du fer dans la cellule (Oudit et al., 2003 ; Mwanjewe et Grover, 2004). Une autre possibilité est de complexer le fer avec la NGAL/24p3 (*Neutrophil Gelatinase-Associated-Lipocalin*), qui est un membre de la superfamille des lipocalines dont le rôle physiologique serait lié à la fixation et au transport de petites molécules. Cette NGAL/24p3 rejoindrait l'endosome et à l'image de la transferrine, sous l'action de l'acidité du pH, elle relarguerait le fer qu'elle transporte (Yang et al., 2002). La voie de transport du fer via la NGAL/24p3 apparaît être importante dans la différenciation des cellules épithéliales durant le développement précoce et les premières phases de l'embryogenèse (Yang et al., 2002).

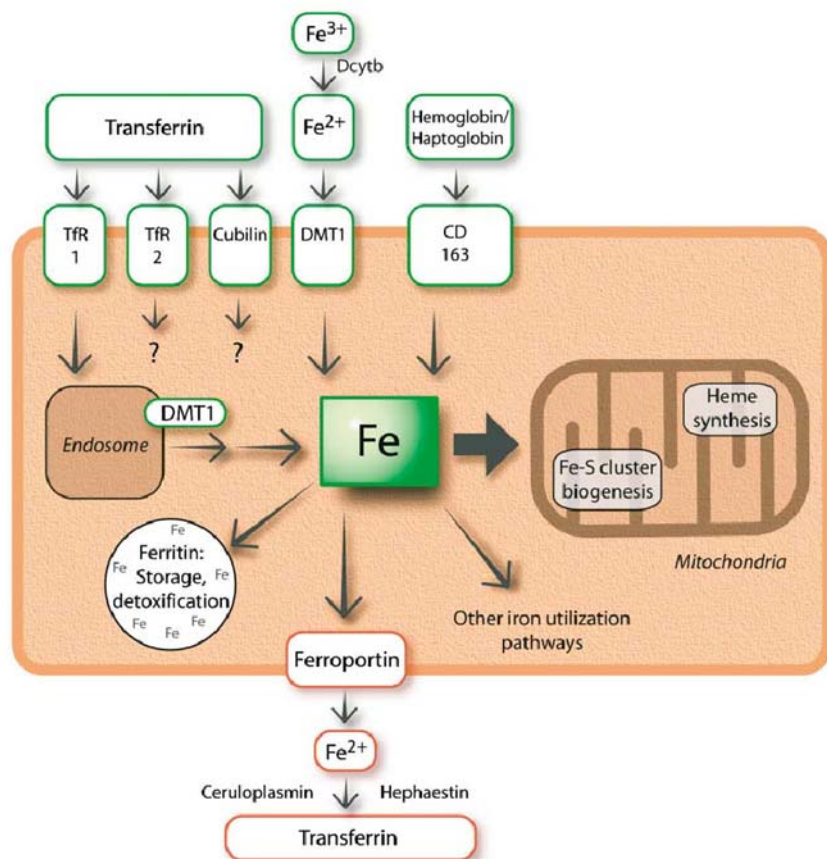
Le fer apporté par l'alimentation est absorbé directement au niveau de la partie apicale des entérocytes. Ce fer est biologiquement indisponible car sous sa forme ferrique et de ce fait insoluble. La première étape consiste donc à le réduire en fer ferreux et l'enzyme responsable de cette opération est la Dcytb qui est une ferrireductase cytochrome « *b-like* » et dont la particularité est d'être une hème-protéine associée avec la membrane apicale des entérocytes (McKie et al., 2001). Le Fe^{2+} peut alors entrer dans la cellule grâce au symporteur DMT-1. DMT-1 a aussi été impliquée dans la recapture du Fe^{2+} dans le rein (Fergusson et al., 2001 ; Hubert et Hentze, 2002) et dans le transport du fer dans le cerveau (Jeong et David, 2003). Le fer peut également être récupéré par les cellules sous d'autres formes complexées à des protéines telles que l'isoferritine (Meyron-Holtz et al., 1999) ou l'hémoglobine. Cette dernière a une importance capitale dans l'hémolyse intravasculaire observée dans certaines maladies comme ce peut être le cas de l'anémie à hématies falciformes qui est liée à la présence de gènes d'hémoglobine anormaux produisant une protéine d'hémoglobine non fonctionnelle, déformant les globules rouges. Dans ces conditions, l'hémoglobine présente dans le plasma est alors capturée par l'haptoglobine, un des marqueurs de l'inflammation, pour être ensuite absorbée par les monocytes et les macrophages exprimant à leurs surfaces des récepteurs spécifiques à cette protéine, les CD163 (Kristiansen et al., 2001).

Ainsi, il existe différents moyens pour la cellule de récupérer du fer, chacun étant adapté aux différentes formes que le fer peut adopter, probablement, afin de permettre une régulation différentielle très fine. Les données biochimiques suggèrent qu'il peut exister des mécanismes

additionnels de récupération du fer mais ils n'ont pas encore été caractérisés au niveau moléculaire ou cellulaire.

La récupération du fer des globules rouges vieillissant ou endommagés est réalisée par une population spécialisée de macrophages qui ne sert qu'au recyclage du fer de l'hémoglobine afin qu'il puisse être réutilisé (Brittenham, 1994). Les globules rouges ainsi phagocytés sont lysés et leur hémoglobine est dégradée grâce à une oxygénase. Une partie du fer récupéré est stocké alors que le reste est exporté et remis en circulation via la transferrine. Ce système est quantitativement important car seulement 1 à 2 mg de fer entre dans le corps par l'intestin chaque jour alors que 25 mg sont nécessaires pour l'érythropoïèse et pour d'autres utilisations. Presque tout ce qui est disponible en fer est donc dérivé de son recyclage par les macrophages (Figure I.2).

Figure I.2 : Métabolisme cellulaire du fer. Représentation d'une cellule générique sur laquelle les voies d'importation et d'exportation du fer sont schématisées. Une grande majorité du trafic intracellulaire du fer est à destination de la mitochondrie où la synthèse de l'hème et de la biogenèse des agrégats Fe-S y sont localisées. D'après Hentze et al. (2004).



1.1.2 Le stockage du fer

Une fois que le fer est entré dans la cellule, la partie qui n'est pas immédiatement utilisée est alors stockée par la ferritine (Ft). La ferritine est une protéine multimérique hautement conservée, qui est constitutivement exprimée (Torti et Torti, 2002). Chez les vertébrés, la ferritine est constituée par 24 chaînes légères (L pour *light*) et autant de chaînes

lourdes (H pour *heavy*). Le complexe multimérique ainsi formé est capable de stocker 4500 atomes de fer. La ferritine séquestre le fer, sous une forme chimiquement moins réactive, la ferrihydrite, ce qui permet à la fois de stocker le fer mais aussi d'avoir une action de détoxification du milieu intracellulaire (Harrison, 1977). Le taux de sous unités H et L varie en fonction du type de tissu, de son statut physiologique mais aussi de sa réponse à une réaction inflammatoire ou infectieuse. La ferritine a comme propriété enzymatique d'être capable de convertir le Fe^{2+} en Fe^{3+} lors de son internalisation. Cette activité ferroxidase est une caractéristique inhérente aux sous unités H. La dégradation de la ferritine est un phénomène concomitant au relargage du fer ce qui permet de remobiliser les stocks de fer dans l'optique de leur prochaine utilisation. Peu de choses sont connues actuellement sur le mode de dégradation de la ferritine mais une chose est quasi certaine, c'est que, aussi bien la voie lysosomale que celle du protéasome sont impliquées dans ce phénomène.

La mitochondrie exprime aussi une ferritine un peu particulière car elle est formée d'un homopolymère de chaîne H, la MtF (*mitochondrial ferritin*) (Levi et al., 2001). La surexpression expérimentale de la MtF entraîne une accumulation de fer dans la mitochondrie diminuant la présence de la ferritine dans le cytosol et augmentant l'expression du TfR1. Il est intéressant de constater que l'expression de la MtF augmente dans les désordres anémiques caractérisés par une accumulation de fer dans la mitochondrie (anémie sidéroblastique), suggérant que la ferritine mitochondriale pourrait aussi jouer le rôle d'un « dépollueur » du cytoplasme préservant ainsi la cellule contre une toxicité trop importante due à une accumulation de fer libre.

1.1.3 L'exportation du fer

Les enterocytes, les macrophages, les hépatocytes, les syncytiotrophoblastes du placenta et les cellules du système nerveux central ont besoin de mécanismes pour pouvoir relâcher du fer de façon très contrôlée afin d'assurer une disponibilité en métal aux endroits où il y en a besoin. Le seul exportateur identifié à ce jour, en tout cas supposé comme tel, est la ferroportine, aussi connu sous le nom de MTP ou de Ireg1 (voir Figure I.2) (Abboud and Haile, 2000 ; Donovan et al., 2000 ; McKie et al., 2000). La ferroportine est localisée sur la membrane basolatérale des enterocytes du duodénum, où elle régule les échanges de fer avec la circulation sanguine apparemment de concert avec l'hephaestine, une ferroxidase homologue à la céruloplasmine, une protéine très abondante dans le plasma (Vulpe et al., 1999).

L'exportation du fer depuis des cellules non intestinales requiert la présence de cette céruloplasmine (Cp) (Harris et al., 1999). La Cp convertit le fer ferreux, vraisemblablement exporté par la ferroportine, en fer ferrique qui est ensuite pris en charge par la transferrine pour son transport dans le plasma (voir Figure I.2). Dans le cerveau, la céruloplasmine, via une ancre glycosylphosphatidylinositol (GPI), est en interaction directe avec la ferroportine (Jeong et David, 2003). Des patients et des souris présentant une déficience au niveau de l'expression de la céruloplasmine, accumulent du fer dans leurs macrophages, leurs hépatocytes et leurs cellules du système nerveux central aboutissant à une érythropoïèse imparfaite et à des problèmes neurodégénératifs. Quelques travaux suggèrent que la Cp pourrait aussi être impliquée dans l'importation du fer dans les cultures cellulaires de tissus (Mukhopadhyay et al., 1998) mais cela reste toujours à confirmer.

1.1.4 La distribution intracellulaire du fer

La quasi-totalité du fer intracellulaire est dirigé vers la mitochondrie, où une quantité importante y est nécessaire pour la biosynthèse de l'hème ainsi que pour la maturation des agrégats fer soufre (voir Figure I.2). Dans les cellules de type érythroïde, où pratiquement tout le fer est utilisé pour la synthèse de l'hème, les endosomes chargés en fer sont supposés former directement des contacts avec la mitochondrie afin de faciliter son approvisionnement en fer (Ponka, 1997). Ce modèle est intrigant, mais pas encore prouvé.

Le fer peut s'accumuler dans la mitochondrie aussi bien chez la levure (Gerber et Lill, 2002) que chez l'homme (Allikmets et al., 1999) lorsque les protéines impliquées dans l'assemblage des agrégats fer soufre sont défectueuses. ABC7, chez l'homme, et son orthologue fonctionnel chez la levure, Atm1p, sont nécessaires à l'exportation des agrégats fer soufre depuis la mitochondrie. Des mutations au niveau du transporteur ABC7 résultent en une anémie sidéroblastique, accompagnée de symptômes neurologiques de type ataxie cérébelleuse (incoordination des mouvements). ABC7 joue aussi un rôle dans la synthèse de l'hème, à travers une interaction directe avec la ferrochélatase, une enzyme qui catalyse la dernière étape de cette voie (Taketani et al., 2003).

Des mutations ou des délétions dans le gène de la frataxine causent des désordres neurodégénératifs plus connus sous le nom d'ataxie de Friedreich. Les patients atteints par ce genre d'ataxie présentent une déficience progressive en protéine contenant des agrégats fer soufre, surtout dans le tissu musculaire et dans les neurones, suivi par une surcharge en fer de plus en plus importante au niveau de leurs mitochondries (Rotig et al., 1997). La frataxine est une protéine mitochondriale localisée dans la matrice. Elle semble être impliquée dans

l'exportation du fer depuis la mitochondrie (Chen et al., 2002) et dans son stockage (Adamec et al., 2000). Des travaux récents sur la levure démontrent clairement que la frataxine est nécessaire pour la biogenèse *de novo* de protéines fer soufre (Muhlenhoff et al., 2003). Elle est retrouvée au centre du complexe responsables de l'assemblage des agrégats fer soufre et serait à l'origine du chargement du fer sur le complexe (Gerber et al., 2003). Il est à noter que l'expression de la frataxine est réduite durant la différenciation des érythroïdes, ce qui pourrait s'expliquer par le fait que toute la machinerie mitochondriale, et de ce fait tous ses stocks de fer, serait employée à la synthèse de l'hème.

Les derniers travaux sur le sujet permettent de comprendre un peu mieux comment le fer rentre dans la mitochondrie et comment sa circulation y est régulée. Récemment, il a été proposé que les protéines fer soufre pourraient agir comme des capteurs de l'état des stocks de fer dans la mitochondrie. Si ces capteurs putatifs viennent à perdre leurs agrégats, le fer mitochondrial s'accumulerait par une incorporation accrue en même temps qu'il y aurait une diminution de son exportation. Cette séquestration du fer pourrait alors continuer jusqu'à atteindre un seuil assez satisfaisant pour qu'il y ait, à nouveau, formation d'agrégat fer soufre. Ensuite, après avoir évalué la quantité de fer restant dans les mitochondries, le pool cytoplasmique pourrait être sondé de la même manière ce qui permettrait de réguler l'absorption du fer au niveau cellulaire (Pandolfo, 2002). Mais les processus de régulation des flux de fer au niveau de la mitochondrie sont plus complexes qu'il n'y paraît et il reste encore beaucoup de mystères à lever à ce niveau là.

1.2 L'homéostasie systémique

L'homéostasie systémique du fer, c'est-à-dire le contrôle de la concentration en fer à travers tout l'organisme, requiert des mécanismes de régulation pour permettre l'entrée ainsi que la mobilisation des réserves emmagasinées. Il n'existe pas, ou tout du moins il n'y a pas encore eu de résultats prouvant le contraire, de voie d'excrétion du fer de l'organisme et de ce fait l'absorption intestinale peut être modulée pour fournir le fer nécessaire au maintien des stocks. Il doit alors y avoir impérativement un système de communication très efficace entre les cellules qui consomment le fer et celles responsables de l'incorporation et du stockage de cet ion métallique (Figure I.3).

1.2.1 Les régulateurs systémiques

Chez un adulte sain, le stockage du fer a lieu dans les hépatocytes ainsi que dans le tissu des macrophages et sa mobilisation n'est effective qu'en réponse à un besoin

précisément identifié. L'absorption intestinale du fer est régulée, en partie, en réponse aux signaux qui communiquent la quantité de fer encore présent dans les réserves via des régulateurs connus sous le nom de « *stores regulator* » (Finch, 1994). Les cellules érythroïdes sont les principales consommatrices de fer et la dotation en fer de la moelle osseuse, des précurseurs des érythroïdes et des globules rouges circulant excède, normalement, la quantité de fer présent dans les réserves. Quand la demande des érythroïdes surpasse les capacités de stockage que les cellules peuvent mettre à disposition de l'érythropoïèse alors l'absorption intestinale est stimulée par un régulateur spécifique de type « *erythroid regulator* » (Finch, 1994). L'homéostasie du fer peut aussi être altérée en réponse à une hypoxie à travers un régulateur humoral dénommé « *hypoxia regulator* » ou suite à un phénomène inflammatoire via un « *inflammatory regulator* ».

L'organisation et la hiérarchie entre les différents régulateurs est fonction de l'état de déficience ou pas en fer durant l'érythropoïèse ainsi que de l'état dans lequel se trouve le niveau des réserves dans les tissus servant à stocker le fer. Les souris hypotransferrinémiques (Trf^{hpx}) présentent un phénotype dans lequel la transferrine est pratiquement absente du plasma (Trenor et al., 2000) ce qui a pour conséquence d'altérer sévèrement l'incorporation du fer au niveau des précurseurs des érythroïdes (Levy et al., 1999) et d'une façon plus générale durant l'érythropoïèse. Cependant, comme nous l'avons déjà vu, les tissus non hématopoïétiques, du fait même du manque de prise en charge du fer par la transferrine, assimilent de façon très importante le fer en utilisant les autres systèmes à leur disposition. Cela a pour conséquence que des organes tels que le foie ou encore le pancréas développent une surcharge anormale en fer. Dans cette situation, le régulateur d'érythroïde (*erythroid regulator*) prend le pas sur le régulateur de stockage (*stores regulator*) ce qui aboutit à une absorption massive de fer via l'intestin et donc à un excès de fer dans les cellules du parenchyme. L'ensemble de ces régulateurs de stockage, d'inflammation, d'hypoxie et d'érythroïde ne sont pas complètement indépendants et permettent une plus grande flexibilité du système de régulation dans lequel chacun pourrait peser de façon quantitativement différente.

Le niveau de fer circulant dans le plasma est déterminé par, à la fois son absorption intestinale et par le recyclage de l'hème via les macrophages (voir Figure I.3). Les effecteurs qui régulent l'absorption intestinale du fer doivent être aussi en mesure de moduler le relargage du fer des macrophages et des hépatocytes. Ceci implique que ces molécules doivent pouvoir agir sur de multiples sites.

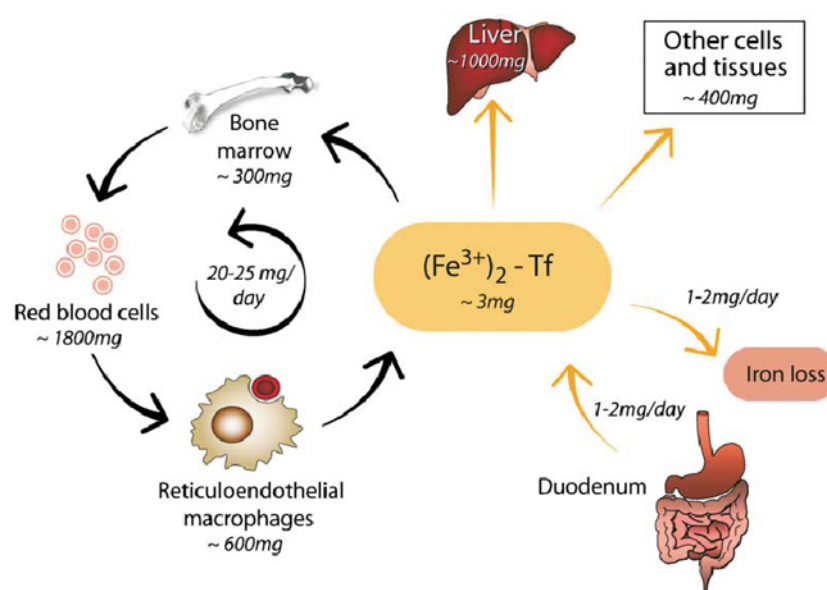


Figure I.3 : Homéostasie systémique du fer. Les principales voies de circulation du fer au sein de l'organisme entre les cellules et les tissus y sont représentées. Les pertes en fer résultent de la desquamation de la peau, des muqueuses et des saignements. Il est important de noter qu'il n'existe pas de voie d'excrétion du fer qui serait contrôlé par l'homéostasie systémique. Les quantités en fer des différents organes et tissus sont relativement stables mais elles peuvent néanmoins être assujetties à des changements d'un individu à l'autre. D'après Hentze et al. (2004).

L'hépcidine (HAMP, LEAP) apparaît être l'un de ces régulateurs. Il s'agit d'un peptide riche en cystéine qui est sécrété par le foie et excrété par le rein et sa ressemblance avec des peptides impliqués dans l'immunité innée laisse à penser qu'il pourrait aussi avoir une activité antimicrobienne (Krause et al., 2000 ; Park et al., 2001 Pigeon et al., 2001). Des souris présentant une perte d'expression de l'hépcidine, ont un niveau de fer très élevé dans le corps sûrement dû à une hyperabsorption du fer associé à une diminution des capacités de stockage des macrophages (Nicolas et al., 2001, 2002a). L'expression de l'hépcidine augmente quand un excès de fer est administré à une souris (Pigeon et al., 2001 ; Muckenthaler et al., 2003), suggérant une limitation de l'absorption intestinale du fer en réponse à ce stimulus alors que des souris transgéniques, exprimant constitutivement l'hépcidine, présentent une carence en fer très marquée (Nicolas et al., 2002a). Ces observations font de l'hépcidine un candidat parfait pour être l'un des régulateurs solubles qui pourraient à la fois agir sur l'absorption intestinale du fer et sur le relargage du fer des macrophages (voir Figure I.4).

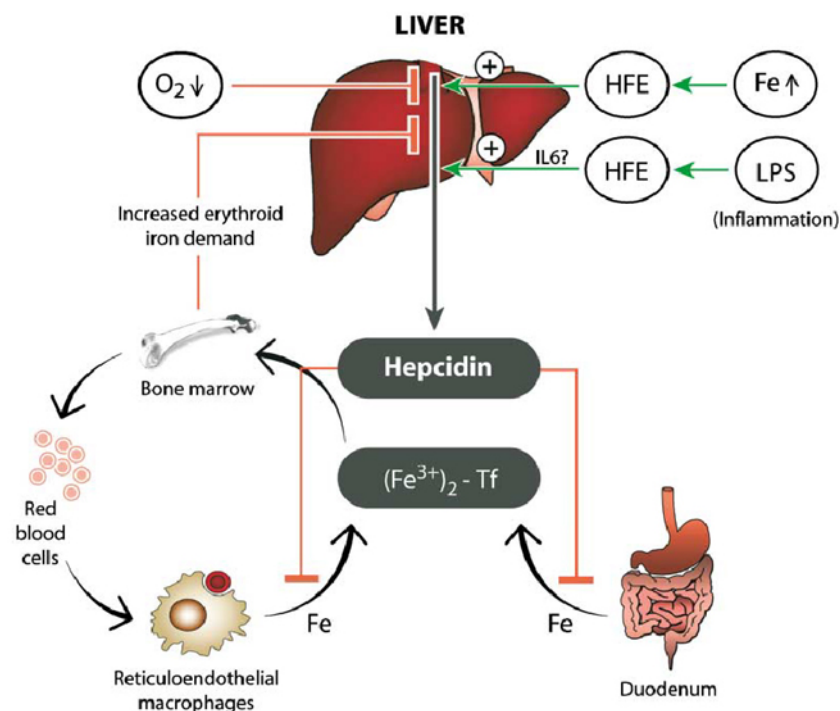


Figure I.4 : Rôle de l'hepcidine, de la protéine HFE et d'autres molécules impliquées dans la régulation de l'homéostasie systémique du fer. L'hepcidine est un peptide sécrété par le foie de type β -défensine. Elle diminue le relargage du fer des macrophages et des entérocytes entraînant par la même occasion une diminution du niveau de fer dans le sérum. L'expression de l'hepcidine est régulée par les niveaux de fer, les stimuli dus aux phénomènes inflammatoires, à la demande en fer des érythroïdes et de l'hypoxie. Des études récentes indiquent que la protéine HFE est nécessaire aux deux voies de stimulation de la synthèse d'hepcidine. D'après Hentze et al. (2004).

L'augmentation du niveau d'expression de l'hepcidine dans un animal auquel une dose importante de fer a été injectée et chez un patient hypertransfusé montre que c'est un élément faisant partie intégrante des régulateurs de stockage. Cependant, une souris Trf^{hpx} , présentant une surcharge en fer, exprime très peu d'hepcidine (Weinstein et al., 2002) suggérant cette fois qu'il s'agirait aussi d'un régulateur des érythroïdes. L'expression de l'hepcidine diminue en réponse à une hypoxie non anémique impliquant cette fois le peptide comme régulateur de l'hypoxie (Nicolas et al., 2002b) alors que son expression est augmentée dans les souris et chez les patients présentant des problèmes inflammatoires ce qui laisse à penser que l'hepcidine jouerait aussi un rôle en tant que régulateur de l'inflammation (Weinstein et al., 2002 ; Nemeth et al., 2003 ; Nicolas et al., 2002b). Bien que ces changements d'expression aient permis de mettre en évidence le fait qu'il existait une corrélation entre l'hepcidine et ces différents régulateurs, aucune voie de signalisation n'a pu encore être établie à ce jour. Il a été prouvé, récemment, que l'interleukine 6 induisait une expression de l'ARNm codant pour l'hepcidine *in vitro* (Nemeth et al., 2003) et *in vivo* (Nemeth et al., 2004) liant ainsi la production d'hepcidine au régulateur de l'inflammation de

façon quasi certaine. Cependant ni le fer ferrique ni l'utilisation d'holotransferrine n'induit l'expression d'hépcidine dans des cultures de cellules hépatiques. Aucun lien moléculaire n'a été trouvé entre l'expression de l'hépcidine et les régulateurs de l'hypoxie et des érythroïdes.

L'hépcidine pourrait être la clé de voûte pour l'ensemble de ces quatre régulateurs. Bien que la production d'hépcidine soit atténuée lorsque l'érythropoïèse vient à manquer de fer, l'activation du régulateur des érythroïdes dans une souris Trf^{hpx} mène à une expression nettement accrue de l'expression des ARNm de DMT-1, de la ferroportine et de Dcytb (McKie et al., 2000, 2001 ; Canonne-Hergaux et al., 2001) L'augmentation de l'expression de ces protéines, relatives au transport du fer, peut résulter en un signal additionnel afin d'augmenter l'absorption de fer en même temps qu'il y a une diminution de la production d'hépcidine.

Aucun récepteur à l'hépcidine n'a été identifié à ce jour mais il est supposé, dans le modèle actuel, qu'elle est capable d'interagir avec une protéine à la surface de la cellule. L'hépcidine, ainsi liée, pourrait entraîner l'inactivation de la machinerie cellulaire d'exportation du fer en diminuant soit l'expression de ses constituants soit en les modifiant. Alternativement, l'hépcidine pourrait lier directement la ferroportine ou une molécule associée directement au transport du fer. La caractérisation des mécanismes d'action de l'hépcidine est, à ce jour, une priorité pour les biologistes. Il est important de résoudre cela afin d'aboutir à une meilleure compréhension du mode de fonctionnement de ce peptide.

1.2.2 Accumulation, invasion et réactivité du fer

Il est généralement accepté que le fer s'accumule dans le cerveau avec l'âge. Cependant, ce processus est tout à fait spécifique à certaines cellules qui permettent l'accumulation du fer dans des régions du cerveau particulièrement visées par des maladies neurodégénératives comme la maladie d'Alzheimer ou la maladie de Parkinson. Les concentrations en fer, non lié à l'hème et la plupart du temps complexé à la ferritine, augmentent dans le putamen, qui correspond à la substance noire de la partie inférieure des hémisphères cérébraux, le cortex moteur, le cortex préfrontal, le cortex sensoriel et le thalamus durant les trente à trente cinq premières années de la vie (Hallgren et Sourander, 1958). Des travaux récents ont montré que les niveaux de H-ferritine, qui est la forme de ferritine préférentiellement exprimé dans les neurones, chez des individus âgés de 67 à 88 ans étaient beaucoup plus élevés dans le cortex frontal, le noyau caudé, le putamen, la substance noire et le *globus pallidus* que chez des individus plus jeunes (27 à 66 ans). Dans le cas de la

L-ferritine, exprimée surtout dans les microglies, cette augmentation n'est visible que dans la substance noire et le *globus pallidus* (voir Figure I.5)(Connor et al., 1995 ; Zecca et al., 2001).

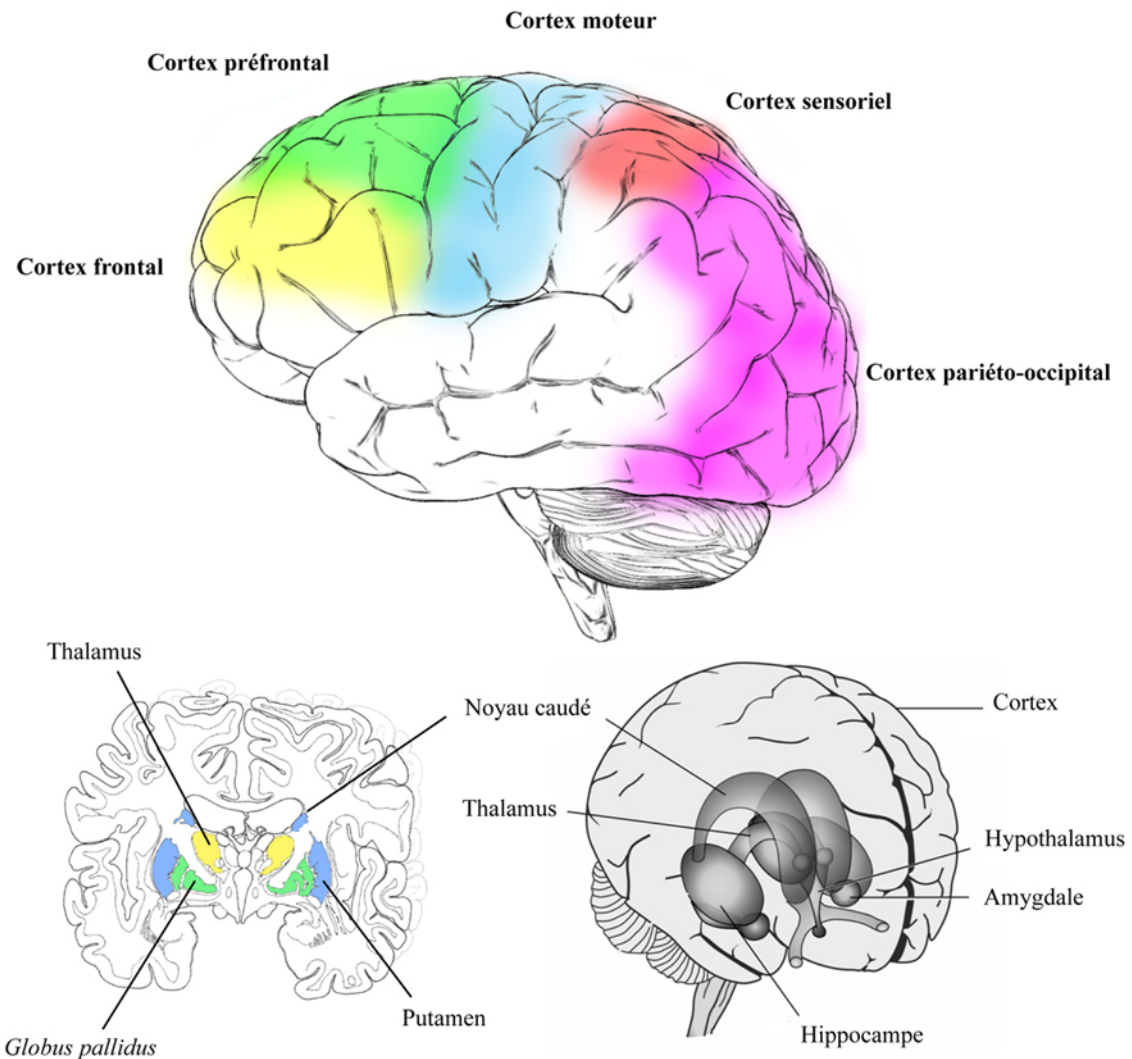


Figure I.5 : Localisation des différentes zones du cerveau présentant des anomalies quant à leur niveau de concentration en fer.

L'expression de l'hème oxygénase 1, responsable de la dégradation de l'hème en bilirubine, augmente aussi avec l'âge. La surexpression de l'hème oxygénase 1 contribuerait à augmenter la sensibilité au stress oxydant des sujets présentant cette accumulation, en exerçant un rôle protecteur en s'opposant aux effets délétères des radicaux libres que l'oxygène génère entre autres au cours de la réaction de Fenton (voir § I.1) (Hirose et al., 2003).

Chez des individus de plus de 80 ans, beaucoup de dépôts de fer sont facilement observables sur les neurones de la substance noire et plus spécialement sur les oligodendrocytes. De la même façon des taux anormalement élevés en fer ont pu être identifiés dans les astrocytes et les microglies du cortex, de l'hippocampe et des amygdales

chez des sujets âgés alors que des sujets jeunes ne présentaient pas ces anomalies. L'accumulation de fer dans les microglies entraînerait ces cellules dans un processus inflammatoire qui contribueraient fortement aux maladies d'Alzheimer et de Parkinson. Toutes ces données ont été collectées à partir de prélèvement *post-mortem* mais des concentrations élevées en fer peuvent être révélées par des méthodes non invasives comme c'est le cas de l'imagerie par résonance magnétique nucléaire (IRM).

L'augmentation des concentrations en fer dans certaines régions du cerveau peut aussi entraîner des altérations dans la vascularisation de ces zones, ce qui est fréquemment observé chez des patients âgés, aboutissant souvent à des maladies neurodégénératives (Brun et Englund, 1986 ; Snowden, 2003 ; Faucheux et al., 1999). L'augmentation des concentrations en fer dans les neurones, les astrocytes et les microglies, qui normalement en contiennent peu, est un élément typique des régions subissant des changements neuropathologiques correspondant aux maladies d'Alzheimer ou de Parkinson comme c'est le cas pour le cortex, l'hippocampe et la substance noire. Cette invasion du fer peut directement endommager les cellules subissant cet occupation, ce qui entraîne une perturbation de l'environnement cellulaire, favorisant la fabrication de toxines et activant des processus de type pathologique comme par exemple des phénomènes inflammatoires, des changements morphologiques ou pire encore l'apoptose. De plus, durant le vieillissement du cerveau, le fer est partiellement converti en hémossidérine et autres dérivés oxyhydroxydes qui contiennent du fer très réactif. Ainsi le prélude au rôle pathogène du fer, dans le vieillissement du cerveau, peut être récapitulé par cette triade : accumulation de fer, invasion et réactivité accrue.

1.2.3 Le fer et les maladies neurodégénératives

1.2.3.1 La maladie de Parkinson

De nombreuses équipes ont montré qu'il existait une corrélation entre l'augmentation totale de la concentration en fer dans les différentes parties du cerveau déjà citées et le niveau de gravité et d'avancement de la maladie (Riederer et al., 1989 ; Dexter et al., 1987 ; Hirsch et al., 1991). Cette augmentation du contenu en fer du cerveau entraîne une modification des voies dopaminergiques via une dégénérescence progressive des neurones dans la maladie de Parkinson (Gerlach et al., 1994). Il y a donc un rapport significatif entre la concentration en dopamine et la concentration en fer. De plus des dépôts en Fe^{3+} ont été trouvés dans la majorité des neurones dopaminergiques des patients ayant atteints un stade avancé de la

maladie, ce qui est tout le temps associé avec une surcharge en ferritine et en neuromélanine (voir Figure I.6) (Jellinger et al., 1990).

Plusieurs expériences ont fourni la preuve que l'accumulation de fer se faisait sous la forme de complexes neuromélanine/fer après un dysfonctionnement dans la gestion des ressources en fer soit dû à une translocation intracellulaire, par exemple du cytosol vers la mitochondrie, soit à une translocation intercellulaire, par exemple des oligodendrocytes vers les neurones (Götz et al., 2004). La concentration en fer, dans le cytoplasme des neurones d'un patient atteint de la maladie de Parkinson, est relativement faible à cause de la forte affinité de la neuromélanine qui joue donc le rôle d'un puissant chélateur. Les complexes neuromélanine/fer activent, *in vitro*, les microglies ou cellules microgliales qui sont des cellules dérivées des monocytes sanguins ayant pénétrées dans le parenchyme du système nerveux central et qui, lorsqu'elles sont activées, se transforment en macrophage entraînant la sécrétion de plusieurs composés neurotoxiques tels que l'interleukine-6, le TNF α pour *tumour necrosis factor α* , l'oxyde nitrique et des radicaux libres. Cette expérience tend à prouver que bien que différents toxiques et autres mécanismes génétiques peuvent initier des dommages neuronaux, la neuromélanine, qui est libérée par les neurones mourants, est potentiellement capable de conduire à l'aggravation de la neurodégénérescence du simple fait qu'elle peut induire la libération des facteurs microgliaux (Wilms et al., 2003 ; Langston et al., 1999).

Une augmentation de l'activité redox du fer associé à la neuromélanine se produit dans les neurones des patients présentant une sévère perte neuronale prouvant ainsi le rôle central de la neuromélanine dans la modulation de la réactivité du fer (Faucheux et al., 2003). La maladie de Parkinson est aussi associée à une augmentation de l'expression de l'hème oxygénase 1 dans les neurones dopaminergiques affectés (Schipper et al., 1998) et à une surexpression du transporteur de cations divalents DCT1 (Qian et al., 1998). Le mécanisme qui est à la base de l'augmentation des concentrations en fer dans le cerveau n'est pas encore connu à ce jour. Il semble être indépendant du système Tf/TfR car il y a une diminution significative de la densité de TfR dans les neurones des patients atteints de la maladie de Parkinson. La principale protéine impliquée dans le stockage du fer, à savoir la ferritine, présente une augmentation significative de sa charge en fer (Griffiths et al., 1999) et plusieurs études ont aussi montré une augmentation de l'expression du récepteur à la lactoferrine chez les patients atteints par la maladie de Parkinson (Faucheux et al., 1995 ; Leveugle et al., 1996). Une autre source potentielle de l'augmentation du niveau de fer serait due à un influx périphérique de fer via le passage à travers la barrière séparant le cerveau de la circulation sanguine. Pour vérifier cette hypothèse Leenders et ses collaborateurs (sous presse) ont utilisé

de la verapamile radiomarquée chez des malades et chez des sujets âgés mais sains. La verapamile est un substrat spécifique de la glycoprotéine P (Pgp). La Pgp est une protéine associée à la membrane et dont la fonction principale est de permettre une détoxification de la cellule via son activité de pompe à efflux. La verapamile n'a pas la possibilité de traverser la barrière hémato-encéphalique et pourtant une forte concentration est retrouvée dans le mésencéphale des patients atteints de la maladie de Parkinson alors que ce n'est pas le cas chez les personnes servant de contrôle. Il semblerait aussi qu'une déficience dans le fonctionnement du système Pgp de la barrière hémato-encéphalique faciliterait l'accessibilité au cerveau du fer présent dans le sérum.

Encore peu de choses sont connues sur les mécanismes pouvant être à la base des dommages cellulaires induits par le fer dans la maladie de Parkinson. Malgré cela, certains acteurs ont quand même pu être mis en lumière. C'est le cas de l' α synucléine qui est à l'origine d'une forme particulière de maladie de Parkinson qui est retrouvée dans certaines familles présentant toutes les mêmes mutations sur le gène exprimant cette protéine. Il a aussi été montré que l' α synucléine est un des principaux composants des Corps de Lewy, qui sont des corps d'inclusion, denses et très riches en fer, présents entre les neurones et qui sont l'un des marqueurs pathologiques de la maladie de Parkinson (voir Figure I.6). Le fer contribue à la synthèse d'espèces réactives de l'oxygène, plus communément appelées ROS pour *Reactive Oxygen Species*, augmentant ainsi le stress oxydant et l'agrégation des protéines au sein même de la cellule (Ostrerova-Golts et al., 2000 ; Uversky et al., 2001 ; Münch et al., 2000). Les études d'Hashimoto (1999) indiquent que les réactions d'oxydation, catalysées par le fer et régulées par le cytochrome c et le peroxyde d'hydrogène, sont fortement impliquées dans l'agrégation de l' α synucléine. En présence de fer et d'agents capables de promouvoir la formation de radicaux libres, tel que la dopamine et le peroxyde d'hydrogène, l' α synucléine produit donc des agrégats intracellulaires qui se distribuent dans tout l'environnement cytoplasmique, interagissant avec les vésicules, les mitochondries et l'ensemble des composants du cytosol ayant pour effet l'activation d'une cascade d'évènements conduisant la cellule vers sa mort. Cependant la question de savoir si l' α synucléine est un neurotoxique ou un neuroprotecteur reste toujours ouverte car il est possible d'interpréter la formation des corps de Lewy et l'accumulation d' α synucléine comme étant un processus compensatoire capable de protéger le neurone de lui-même.

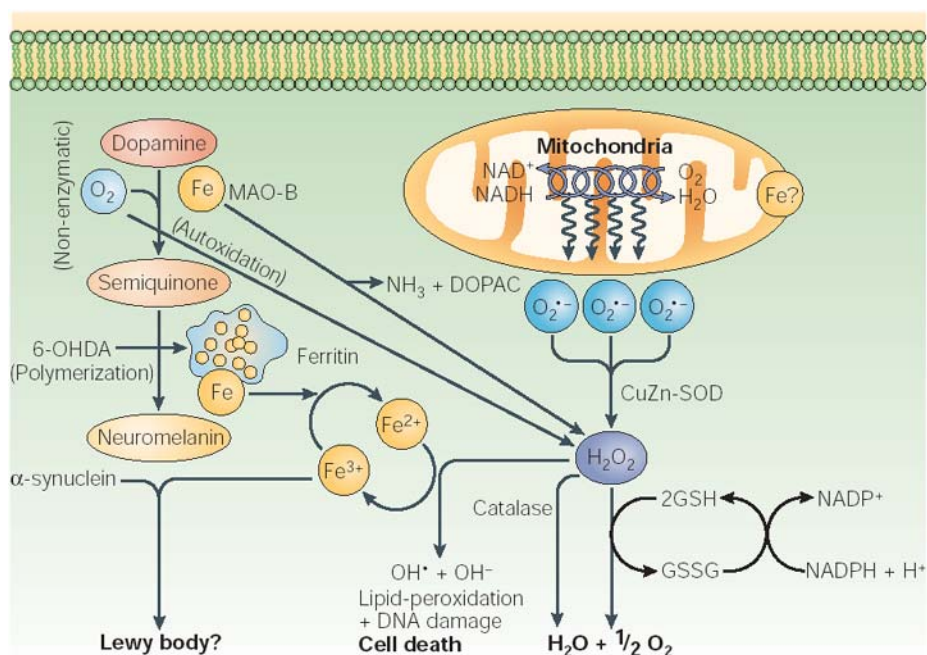


Figure I.6 : Schéma présentant les mécanismes à la base de la maladie de Parkinson. Ce modèle essaye d'expliquer les différentes cascades pouvant conduire la cellule à la mort, il est basé sur des observations réalisées *post-mortem*. Dans l'ensemble des cas il y a une réduction de l'activité du complexe mitochondrial I, une perte de glutathion réduit (GSH), une augmentation de la concentration en fer, une augmentation des marqueurs du stress oxydant, une augmentation de la rapidité de renouvellement de la dopamine ou voir carrément de sa disparition, une apparition d' α synucléine pathologique, une génération de corps de Lewy et une dégénérescence membranaire des neurones touchés. Des mécanismes similaires impliquant le stress oxydant induit par le fer et le peroxyde d'hydrogène ont pu être observés dans d'autres maladies neurodégénératives. CuZn-SOD, superoxyde dismutase à cuivre et à zinc ; GSSG, glutathion oxydé ; H_2O_2 , peroxyde d'hydrogène ; MAO-B, monoamine oxydase B ; NAD, nicotinamide adénine dinucléotide ; OH^\bullet , radical hydroxyle ; $O_2^{\bullet-}$, anion radicalaire superoxyde ; OH^- , anions hydroxydes ; 6-OHDA, 6-hydroxydopamine. D'après Zecca et al. (2004).

1.2.3.2 La maladie d'Alzheimer

L'accumulation de fer dans le cerveau, particulièrement dans les cellules qui sont associées aux plaques séniles, elles mêmes constituées des plaques amyloïdes, est une observation constitutive à la maladie d'Alzheimer et est extrêmement étudiée (Levine et al., 2004). Dans le cerveau des patients atteints de cette maladie, l'accumulation de fer se produit en plus de l'augmentation normale et relative au vieillissement, ce qui a pour conséquence une augmentation accrue des risques dûs au stress oxydant (Connor et al., 1992 ; Thompson et al., 2003). Des résultats récents semblent indiquer que la sénescence des cellules et la maladie d'Alzheimer sont toutes les deux associées à un déclin en myéline des neurones ce qui serait le résultat d'un contenu anormalement élevé en fer de cette même gaine (Bartzokis, 2004). Le fer a donc un impact direct sur la formation des plaques amyloïdes à travers ses effets sur la production des précurseurs de ces plaques, les APP pour *Amyloid Precursor Protein* (Rogers et al., 2002). D'autre part, la capacité de l' α -sécrétase à cliver l'APP peut, elle aussi, être modulée par le fer (Bodovitz et al., 1995). Le fer semble donc promouvoir aussi bien le dépôt

de l'amyloïde β sous forme de plaques que le stress oxydant (Huang et al., 2000 ; Rottkamp et al., 2001 ; Mantyh et al., 1993). Il existe actuellement un débat dans la communauté scientifique sur le rôle protecteur de l'amyloïde β qui, sous sa forme complexée, serait une sorte d'isolant censé protéger l'environnement des neurones sains du stress oxydant des neurones malades mais ce n'est pas encore une idée complètement reconnue (Perry et al., 2002). Aussi, une meilleure compréhension des mécanismes cellulaires impliqués dans la régulation du fer au niveau du cerveau est fondamentalement importante afin de permettre l'identification de toutes les bases biologiques de la maladie d'Alzheimer.

Un lien entre la surcharge congénitale en fer, hémochromatose, et la maladie d'Alzheimer a récemment été mis à jour. La protéine HFE, à l'origine de l'hémochromatose, est localisée dans les vaisseaux sanguins circulant dans le cerveau ainsi que dans les cellules associées aux plaques séniles (Connor et al., 2001). La protéine HFE est aussi exprimée dans les astrocytes des patients atteints de la maladie d'Alzheimer ainsi que dans leurs neurones. De la même façon, il est possible de colocaliser la protéine TAU, une protéine qui normalement se lie aux microtubules afin de promouvoir leur assemblage et leur stabilité mais qui, dans le cas de la maladie d'Alzheimer, se retrouve localisée avec les microfibrilles faisant d'elle un des premiers marqueurs de la maladie, avec la protéine HFE. La présence de protéine HFE mutée, chez les patients atteints d'Alzheimer, appuie fortement l'idée que la dérégulation de la balance gérant les niveaux de concentration en fer dans le cerveau contribue à la progression de la maladie et sa prédominance laisse à penser que cela pourrait même être un facteur important de risque (Moalem et al., 2000). Les protéines HFE mutées sont associées à une augmentation du stress oxydant et par la même occasion au niveau de gravité de la maladie comme cela a pu être évalué neuropathologiquement (Braak et al., 1993 ; Pulliam et al., 2003). Cela fournit un indice supplémentaire au fait que le fer, à la base de l'augmentation du stress oxydant, favorise la neurodégénérescence au cours de la maladie d'Alzheimer.

Le sous type C2 de la transferrine a aussi été trouvé à des taux anormalement élevés chez des patients atteints d'Alzheimer. La présence du variant C2, en plus d'un mutant HFE, augmente le risque de développer la maladie d'un facteur cinq (Zambenedetti et al., 2003 ; Hussain et al., 2002 ; Robson et al., 2004). Il n'existe pas de différences significatives entre la TfC2 et la Tf dans leur façon de fixer le fer (Van Landeghem et al., 1998) et donc l'impact de la mutation de TfC2 sur la maladie d'Alzheimer n'est probablement pas due à sa capacité à délivrer le fer, d'une manière différente de celle de la transferrine, au niveau du cerveau et des neurones. Des données combinées sur les mutations de HFE et de la Tf ont clairement indiqué

que les altérations génétiques, affectant spécifiquement les protéines impliquées dans le contrôle de la concentration du fer, pouvaient augmenter les risques de maladie et qu'elles fournissaient une très forte preuve que la dérégulation des niveaux de fer dans le cerveau contribuait à l'apparition de la maladie d'Alzheimer.

1.2.3.3 Les autres maladies neurologiques

Le fer, et plus particulièrement son accumulation excessive, a été impliqué dans une série d'autres maladies neurologiques (Ke et al., 2003) et quelques unes d'entre elles sont brièvement discutées ici.

Le fer est intimement lié à la Cp pour ce qui est de son exportation depuis les cellules vers la circulation sanguine (voir § I.1.1.3). Cependant dans l'aceruloplasminémie congénitale, qui est caractérisée par une neurodégénérescence progressive du système nerveux central ainsi que par une ataxie, des mutations dans le gène codant pour la Cp causent une absence totale de la protéine dans le plasma (Yoshida et al., 1995 ; Harris et al., 1995 ; Miyajama et al., 1997). Cet état de fait entraîne un sévère excès de fer au sein des tissus du parenchyme y compris dans le cerveau et plus particulièrement dans les ganglions de la base qui regroupent le noyau caudé, le putamen et le *globus pallidus*. Les implications sont que la Cp est nécessaire pour la mobilisation du fer depuis les cellules (Harris et al., 1999). L'ataxie de Friedreich est la plus commune des ataxies et elle est caractérisée par une dégénérescence d'une grande partie des neurones sensoriels, de l'appareil spinocérébrale et par une cardiomyopathie (Patel et Isaya, 2001). La maladie est causée par une réduction substantielle de la concentration en frataxine, une protéine mitochondriale (voir § I.1.1.4), qui est provoqué par l'insertion d'un triplet répété GAA dans le premier intron du gène aboutissant à une réduction de l'expression de la frataxine par inhibition de la transcription (Zoghbi et Orr, 2000). Cela mène à une augmentation en fer du contenu de la mitochondrie, qui semble refléter le rôle vital de la frataxine dans la biosynthèse des agrégats fer soufre (Gordon, 2000 ; Muhlenhoff et al., 2002).

La neuroferritinopathie est une maladie héréditaire qui touche aussi les ganglions de la base, présent dans la zone extrapyramidale du cerveau, et qui n'est causée que par la simple insertion d'une adénine dans le gène codant pour la chaîne L de la ferritine. Ceci entraîne une altération de la partie carboxy-terminale de la protéine (Curtis et al., 2001). Les patients affectés présentent un dépôt de fer dans les ganglions de la base, une agrégation anormale de la ferritine, du fer dans le *globus pallidus* et dans la substance noire ainsi qu'une faible

concentration en ferritine (Crompton et al., 2002). Une autre mutation, avec l'insertion de deux bases dans l'exon 4 du gène de la chaîne L de la ferritine, causant un changement dans la séquence primaire de la protéine ainsi que dans la longueur de la chaîne polypeptidique, a aussi été décrite. Les patients atteints de neuroferritinopathie ont une accumulation anormale de ferritine dans les neurones et dans les cellules gliales ainsi qu'une sévère perte neuronale. Cliniquement, ils sont affectés par une incoordination de leurs mouvements et par un déclin progressif de leurs fonctions cognitives (Vidal et al., 2004).

Des mutations dans le gène qui code pour la pantothénate kinase 2 (PANK2) ont été montrées comme étant le principal défaut génétique associé à une neurodégénérescence avec une accumulation de fer dans le cerveau, plus communément connu sous le nom de syndrome de Hallervorden-Spatz (Zhou et al., 2001 ; Hayflick, 2003). Cette maladie autosomale récessive est caractérisée par une dystonie, qui est une contraction involontaire et douloureuse d'une ou plusieurs parties du corps, une rétinopathie pigmentaire chez l'enfant et une déficience neuropsychiatrique de la parole chez l'adulte. La panthoténate kinase est nécessaire pour la biosynthèse du coenzyme A et son adressage vers la mitochondrie. Le modèle actuel propose que la cystéine, qui se condense avec le phosphopanthoténate pour donner du coenzyme A, s'accumulerait entraînant ainsi une chélation du fer causant un important stress oxydant. Ceci aboutirait à l'augmentation de la concentration en fer dans les ganglions de la base avec toutes les conséquences néfastes que cela induit.

Il semble donc clair qu'une mauvaise gestion du fer au niveau du cerveau, mais d'une façon plus générale au niveau de l'ensemble de l'organisme, peut avoir des conséquences désastreuses aboutissant à une issue fatale, c'est ce qui place le fer au centre de la recherche sur les maladies neurodégénératives.

2 Le système IRP/IRE

Les protéines impliquées dans l'incorporation, le stockage, l'utilisation et l'exportation du fer doivent donc être régulées de façon très stricte et d'une manière très coordonnée. Les signaux et les mécanismes, qui orchestrent leur expression, impliquent une modulation de leur transcription que ce soit en l'empêchant ou en modifiant la stabilité de leurs ARNm. Le système auquel est dévolu ce rôle est le système IRP/IRE. Les protéines régulatrices du fer, les IRP pour *Iron Regulatory Proteins*, interagissent avec les IRE (*Iron Responsive Elements*) qui sont des structures conservées en forme d'épingles à cheveux trouvées dans les portions

non traduites, les UTR pour *UnTranslated Region*, des ARNm codants pour les protéines relatives à la gestion de l'homéostasie du fer.

2.1 IRP1 et IRP2

Il existe deux protéines IRP, IRP1 et IRP2 qui sont toutes les deux des homologues cytoplasmiques respectivement de 889 et 964 acides aminés et elles appartiennent à la famille des isomérases à agrégat fer soufre (Eisenstein, 2000 ; Pantopoulos et Hentze, 2000 ; Cairo et Pietrangelo, 2000). La protéine humaine IRP1 partage 61% d'identité de séquence et 80% d'homologie avec la protéine IRP2 humaine. De plus, IRP1 est à 22% identique à l'aconitase mitochondriale de porc, un membre bien caractérisé de la famille des isomérases à agrégat fer soufre. Par analogie à la structure connue de l'aconitase mitochondriale, IRP1 et IRP2 ont été subdivisées en trois domaines compacts liés à un quatrième par une région charnière flexible (Robbins et Stout, 1989 ; Lauble et al., 1992) (Figure I.7). Une différence notable existe malgré tout entre IRP1 et IRP2. En effet, IRP2 contient une insertion, riche en région cystéine et en région proline, de 73 acides aminés imbriquée dans le domaine 1.

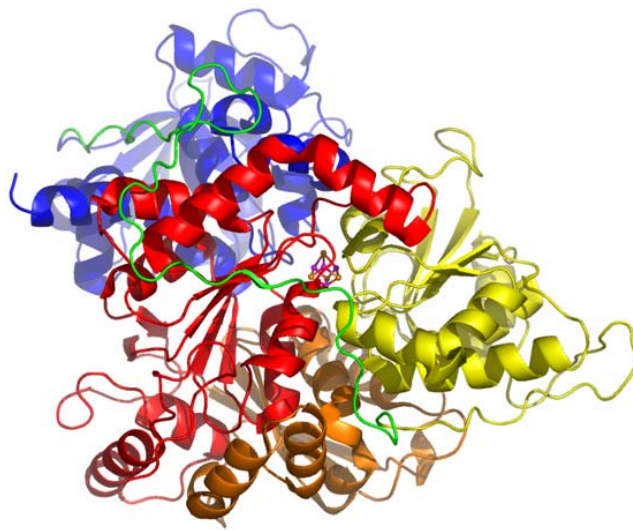


Figure I.7 : Structure de l'aconitase mitochondriale de bœuf (code pdb 1C96). Les couleurs représentent les différents domaines de la protéine : le domaine 1 en rouge, le domaine 2 en orange, le domaine 3 en jaune, le linker en vert et le domaine 4 en bleu.

IRP1 est exprimé de façon constitutive alors qu'il a été initialement proposé que IRP2 n'était produit que de façon spécifique dans certains tissus mais il apparaît maintenant que l'abondance d'IRP2 dans les lignées cellulaires et les tissus avait été prématurément sous-estimée (Rouault, 2002). Peu de choses sont connues sur la contribution relative de chacune des deux IRP dans le maintien de l'homéostasie cellulaire du fer. Des expériences *in vitro* ont montré qu'IRP1 et IRP2 ne différaient pas dans leur capacité à réguler la traduction de la ferritine (Kim et al., 1995), ni dans celle de lier les ARNm de la Ft et du TfR. Cependant dans

certaines lignées de cellules IRP2 semble jouer un rôle prépondérant dans le contrôle de la régulation du TfR et de l'expression de la Ft (Recalcati et al., 1998 ; Kim et Ponka, 1999 ; Kim et Ponka, 2002). D'un autre côté, l'expression d'un mutant constitutif d'IRP1 (C437S) dans des cellules RD4 humaines ou des cellules H1299, ces deux lignées étant des cellules cancéreuses, est suffisant pour bouleverser la régulation par le fer de la Ft et du TfR (DeRusso et al., 1995 ; Wang et Pantopoulos, 2002). Il est concevable d'imaginer que les capacités de régulation de chacune des protéines IRP puissent être modulées par des facteurs cellulaires spécifiques. L'étude de cellules dérivées d'animaux IRP1^{-/-} et IRP2^{-/-}, est attendue pour lever le voile sur cette question.

Des souris « *knock-out* » IRP1^{-/-} et IRP2^{-/-} ont été générées dans le laboratoire de Tracey A. Rouault au NIH à Bethesda et les perturbations engendrées par IRP1^{-/-} ne produisent pas de phénotypes clairement marqués (Rouault, 2002) alors que ce n'est pas le cas pour les souris IRP2^{-/-} qui montrent une régulation de homéostasie du fer complètement aberrante, accumulant du fer dans la muqueuse intestinale et dans le système nerveux central (LaVaute et al., 2001, Meyron-Holtz et al., 2004). Ces données établissent qu'IRP2 est un important régulateur du métabolisme du fer.

2.1.1 IRP1

IRP1 est régulée par le comportement peu habituel de son agrégat fer soufre (Haile et al., 1992a ; 1992b). Dans une cellule pourvue d'un stock suffisant en fer soluble, IRP1 assemble un agrégat cubique [4Fe-4S] qui inhibe son activité de fixation aux IRE. De plus, cet agrégat convertit la protéine IRP1 en une aconitase cytosolique qui possède la même efficacité catalytique que son homologue mitochondriale (Emery-Goodman et al., 1993 ; Kennedy et al., 1992 ; Brazzolotto et al., 1999). Cette activité d'isomérase permet à IRP1 de convertir le citrate en isocitrate. Cette isocitrate peut à son tour être transformé en α cétooglutarate par l'isocitrate déshydrogénase. Au cours de cette réaction le NADPH produit peut servir à réduire la forme disulfide du glutathion (GSSG) en forme sulfhydryle (GS) grâce à l'intervention de la glutathion réductase (Kim et al, 2004). La forme réduite du glutathion intervient alors dans la détoxification de la cellule en réagissant avec le peroxyde d'hydrogène et les peroxydes organiques pour donner de l'eau, un dérivé alcool et de la GSSG. IRP1 joue donc un rôle dans le contrôle des activités redox de la cellule en stimulant la synthèse de NADPH cytoplasmique (Narahari et al., 2000). Ainsi, IRP1 est une protéine bifonctionnelle dans laquelle chacune des activités est mutuellement exclusive (Figure I.8). En l'absence de fer dans la cellule, l'agrégat se désassemble par un mécanisme lent (Pantopoulos et al., 1996)

jusqu'à ce que l'apoIRP1 acquière la capacité de fixer l'ARN. C'est le résultat d'un réarrangement structural qui a pour conséquence le remaniement de la gorge, accueillant l'agrégat, afin de pouvoir permettre les interactions avec l'IRE (Gegout et al., 1999 ; Brazzolotto et al., 2002).

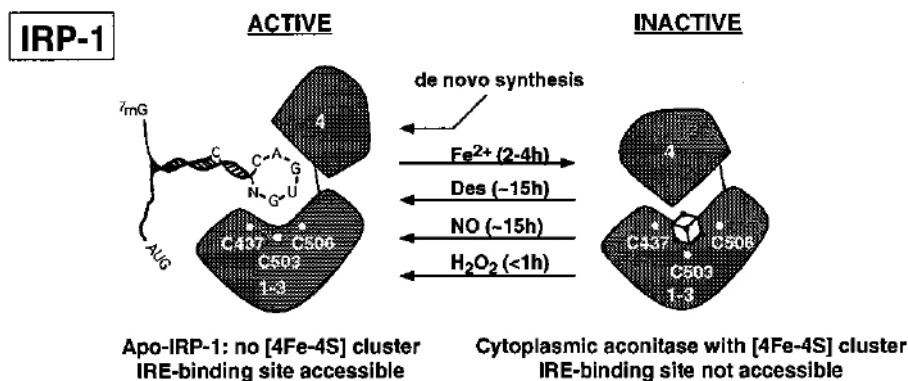


Figure I.8 : Régulation d'IRP1. L'assemblage et le désassemblage de l'agrégat fer soufre est le mode de régulation qui permet à IRP1 de passer d'une forme active, c'est-à-dire capable de lier l'ARN, à une forme inactive possédant une activité aconitase et ce en réponse à différents stimuli. D'après Hentze et Kühn (1996).

La topologie de l'agrégat a été cartographiée par des analyses de mutagenèse, ce qui a permis d'identifier trois cystéines (C437, C503 et C506) liant les fers de l'agrégat au squelette de la protéine alors que le quatrième fer lie une molécule d'eau du solvant et interagit avec le citrate qui est le substrat de la forme aconitase de la protéine IRP1. De même, la sérine 778 semble indispensable à l'activité aconitase de la protéine car son remplacement l'inactive malgré la présence d'un l'agrégat fer soufre correctement assemblé (Philpott et al., 1994). Un dernier acide aminé semble important à l'activité isomérase d'IRP1, il s'agit de l'arginine 699, qui, selon à la structure de l'aconitase mitochondriale, permet la fixation du substrat et sa mutation dans IRP1 entraîne une perte d'activité (Hirling et al., 1994).

Les modifications des résidus responsables de la coordination du centre [4Fe-4S], notamment celles résultant des mutations conduisant aux formes C437S et C503S/C506S d'IRP1 (Hirling et al., 1994 ; Philpott et al., 1994), génèrent des protéines constitutivement capable de fixer le motif IRE. De plus, l'état réduit de ces cystéines est une condition importante pour la fixation de l'ARN (Philpott et al., 1994). Les résidus lysines et arginines des protéines sont souvent importants dans la liaison aux acides nucléiques (Puglisi et al., 1992 ; Harada et Frankel, 1995). Ainsi, la substitution des arginines 536, 541 ou 780 diminue l'affinité d'IRP1 pour le motif IRE (Philpott et al., 1994). Grâce à des expériences de réticulation du motif IRE sur IRP1, un peptide en interaction avec l'ARN, correspondant aux résidus 121 à 130, a pu être identifié (Basilion et al., 1994). Plus récemment, une étude de mutagenèse dirigée contre divers peptides, supposés être en surface de la protéine, a mis en

évidence l'importance des résidus d'arginine 728 et 732 dans le contact avec l'IRE, notamment au niveau de la cytosine non appariée, et la région 685-689 dans la reconnaissance de la boucle d'ARN (voir § I.2.2) (Kaldy et al., 1999). Il est aussi à noter que des mutations, permettant de mimer l'état phosphorylé de la sérine 138 (S138D ou S138E), affectent négativement l'assemblage de l'agrégat (Brown et al., 1998) et que IRP1, ainsi sensibilisée, est plus facilement dégradée en présence de fer (Fillebeen et al., 2003). Ceci suggère une régulation additionnelle au niveau de la stabilité de la protéine. Des expériences similaires ont prouvé que la sérine 711 est un résidu critique au contrôle de l'activité IRP1. En effet, un mutant S711E présente des problèmes aussi bien en ce qui concerne sa capacité à fixer les motifs IRE que dans son activité de conversion du citrate en isocitrate (Fillebeen et al., 2005).

2.1.1.1 Les protéines fer-soufre

Dès 1960, les protéines à agrégat fer-soufre ont été identifiées comme étant des composants essentiels au développement des organismes vivants. A cause de leur structure et de leur propriétés chimiques (Beinert et al., 1997 ; Johnson, 1998), les agrégats fer-soufre sont des agents idéaux pour accepter, donner, transférer ou stocker des électrons. En effet, la combinaison chimique du fer et du soufre permet à ces agrégats d'avoir des potentiels redox dans la majorité des cas très bas, de l'ordre de -600 à -150 mV, ce qui favorise leur utilisation, par la cellule, en temps qu'agent oxydant ou reductant. Pendant bon nombre d'années les scientifiques ont pensé que le seul intérêt de ces agrégats résidait dans leur pouvoir d'oxydoréduction mais petit à petit, les découvertes aidant, les propriétés des agrégats Fe-S se sont multipliées et de nouvelles fonctions ont commencé à émerger. Il n'est pas rare maintenant de trouver ces agrégats fer-soufre impliqués dans le transfert d'électrons comme c'est le cas dans l'hydrogénase à [NiFe] de *D. gigas* (Volbeda et al., 1995), où trois agrégats agencés de façon quasi linéaire permettent d'évacuer, du site actif, les électrons produits durant la catalyse vers un accepteur naturel. Il est aussi possible de trouver des agrégats capable de fixer et d'activer des substrats comme ce peut être le cas de l'aconitase mitochondriale (Beinert et al., 1996) ou encore capable de stabiliser des radicaux (Staples et al., 1996). Les agrégats peuvent aussi fonctionner en couplant le transfert d'électron à un transport de proton (Duff et al., 1996) et dans les protéines « Rieske », l'ionisation des groupes imidazoles peuvent être couplées, eux, à une réaction d'oxydoréduction (Iwasaki et al., 1996). L'agrégat peut aussi jouer un rôle structural, en fixant des ligands cystéines de différentes sous-unités, il peut permettre la formation d'un dimère comme c'est le cas de la nitrogénase (Howard et Rees, 1996) et l'agrégat F_x du photosystème I (Golbeck et Bryant,

1991). Les agrégats fer soufre sont aussi capables de stabiliser les structures de certaines protéines aux fonctions un peu spéciales comme c'est le cas de l'endonucléase III de *E. coli* dans laquelle un site de fixation un peu particulier n'est possible que grâce à la présence de l'agrégat (Thayer et al., 1995). Toujours d'un point de vue structural, des agrégats ont été montrés comme étant capables de protéger leurs protéines de l'attaque des protéases intracellulaires comme dans l'amidotransférase de *Bacillus subtilis* (Grandoni et al., 1989).

Plus intrigant c'est la découverte d'agrégat Fe-S capable de jouer le rôle de capteur de fer, d'oxygène, d'anion superoxyde ou encore d'oxyde nitrique (Rouault et al., 1992 ; Hentze et Kühn, 1996 ; Beinert et Kiley, 1996 ; Gaudu et Weiss, 1996 ; Hidalgo et al., 1995). Cela semble se produire selon deux modes distincts. Dans le premier, l'oxydation de l'agrégat par l'oxygène peut fournir le signal de l'activation des mécanismes de défense contre l'anion superoxyde comme cela a pu être observé avec la protéine SoxR de *E. coli* (Gaudu et Weiss, 1996 ; Hidalgo et al., 1995) alors que dans le second l'état d'oxydation de l'environnement induit l'assemblage ou le désassemblage de l'agrégat fournissant ainsi le signal. C'est le cas du dimère de FNR, chez *E. coli*, qui est actif jusqu'à sa dissociation en monomère suite à la destruction de son agrégat fer soufre (Khoroshilova et al., 1995).

C'est en cela que le cluster d'IRP1 est peu habituel car la fonction même de la protéine est modulée par ce jeu d'assemblage et de désassemblage de son agrégat [4Fe-4S]. Au-delà du rôle de capteur de fer que peut revêtir ce mécanisme peu commun, la perte de la fonction aconitase entraîne l'acquisition d'une nouvelle fonction par la protéine, à savoir sa capacité à se fixer à de l'ARN. Ce mode de régulation par désassemblage de l'agrégat fer soufre dépasse de loin la simple inactivation de la protéine car contrairement à tout ce qui avait été décrit jusqu'à présent IRP1 ne devient pas inactive, comme la FNR, mais au contraire elle voit son activité changer du tout au tout et c'est ce qui la rend d'autant plus intéressante à étudier.

2.1.1.2 Les homologues de IRP1

Depuis ces dernières années les membres de la famille des aconitases ont particulièrement bien été étudiés mais depuis peu quelques nouvelles fonctions sont mises à jour, des fonctions qui, par le passé, n'étaient dévolues qu'aux simples protéines IRP. Ainsi la protéine AcnB de *E. coli*, qui est aussi un membre de cette grande famille, est capable d'acquérir une activité de régulateur transcriptionnel en fonction de l'état de son agrégat [4Fe-4S] (Cunningham et al., 1997). Au même titre qu'IRP1, elle possède une activité aconitase cytoplasmique qui laisse place à une activité de fixation de l'ARN sur certains motifs proches structuralement des motifs IRE des vertébrés mais avec néanmoins une affinité plus faible que

celle d'IRP1 pour les IRE (Tang et Guest, 1999). Le mode de régulation est fonction de la disponibilité du fer dans la cellule et des conditions de stress oxydant y régnant. Récemment la structure d'AcnB a été résolue (Williams et al., 2002) et malgré la faible identité existante entre les séquences de AcnB et celle de l'aconitase mitochondriale de bœuf, il y a un haut degré de conservation structurale entre les deux protéines avec seulement quelques différences majeures et quelques réarrangements du fait même de l'existence d'un domaine supplémentaire dans AcnB. Ceci semble être en accord avec le fait que la famille des aconitases pourrait être séparée en trois sous-familles structurales très proches comme cela avait été proposé par Gruer et ses collaborateurs (1997).

Une nouvelle aconitase Aco1p vient d'être caractérisée chez la levure comme étant potentiellement capable à son tour de fixer non pas de l'ARN mais de l'ADN mitochondrial après que son agrégat [4Fe-4S] eut été désassemblé (Chen et al., 2005). Il semblerait que ce soit un moyen pour la cellule de préserver son ADN mitochondrial, indispensable au maintien de la chaîne respiratoire, dans des conditions de stress oxydant.

2.1.2 IRP2

IRP2 est régulée par la nécessité d'approvisionner la cellule en fer et en oxygène. Dans ce but, elle est synthétisée *de novo* lorsque le manque est détecté (Henderson et Kühn, 1995 ; Hanson et al., 1999), restant stable le temps nécessaire au retour à la normale et elle est ensuite dégradée via le protéasome (Iwai et al., 1995). Le modèle actuellement reconnu comme étant responsable de la dégradation d'IRP2 implique l'insertion d'un domaine dit « de dégradation » de 73 acides aminés (Iwai et al., 1998) qui aurait pour fonction d'être reconnu par le protéasome après l'oxydation spécifique de trois résidus cystéines (C168, C174 et C178) (Figure I.9). Cette oxydation pourrait faire suite à la fixation directe d'un fer sur ces cystéines (Iwai et al., 1998). Des expériences, *in vitro*, menées sur des peptides recombinants portant ce type de domaine de dégradation ont montré une oxydation de leurs cystéines en fonction de la présence ou non de fer soluble dans le milieu (Kang et al., 2003). De plus une étude de double hybride, avec comme sonde le domaine de dégradation, a permis l'identification d'une protéine possédant une activité E3 ubiquitine ligase appelé HOIL-1 (*Heme-oxidized IRP2 ubiquitin ligase-1*) (Yamanaka et al., 2003). Le problème réside dans le fait que ce travail propose que le signal qui permet la reconnaissance d'IRP2 par la HOIL-1 n'est pas l'oxydation des cystéines mais plutôt la fixation d'un hème à la protéine, ce qui avait déjà été proposé par le passé (Goessling et al., 1998). Des résultats récents montrent qu'un hème est capable de se lier à IRP2. Cette hème, fixé à la protéine, réagirait avec l'oxygène

moléculaire présent dans la cellule pour oxyder les cystéines en éliminant les soufres convertissant ces cystéines en alanines. Cette modification covalente pourrait être le signal moléculaire pour la dégradation d'IRP2 (Jeong et al., 2004). Cette théorie est néanmoins incompatible avec la fixation d'un fer par les trois cystéines du domaine de dégradation. Il est aussi clair que le domaine de dégradation de 73 acides aminés n'est pas indispensable à la protéolyse d'IRP2. Ainsi un triple mutant C168S, C174S et C178S, ou même une protéine IRP2 dépourvue de ce domaine de dégradation, présente toujours la même sensibilité à la dégradation en fonction de la concentration en fer (Wang et al., 2004 ; Hanson et al., 2003). Ces résultats suggèrent fortement que le modèle d'oxydation des cystéines n'est pas encore complètement satisfaisant pour expliquer la dégradation de la protéine IRP2. Il reste donc encore pas mal de travail à effectuer pour arriver à comprendre le mécanisme de dégradation d'IRP2.

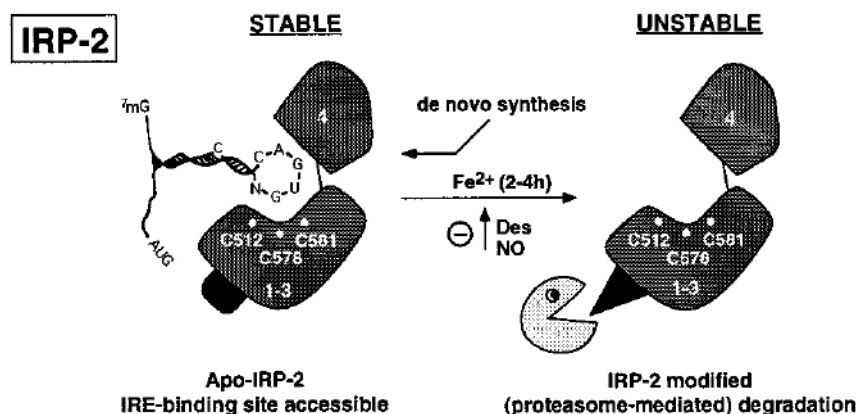


Figure I.9 : Régulation d'IRP2. La régulation passe, contrairement à IRP1, par une étape de dégradation par le protéasome de la protéine. Malgré la conservation des cystéines (C512, C578 et C581), IRP2 ne semble pas être en mesure d'accueillir un agrégat fer soufre. D'après Hentze et Kühn (1996).

2.2 Les cibles des IRP

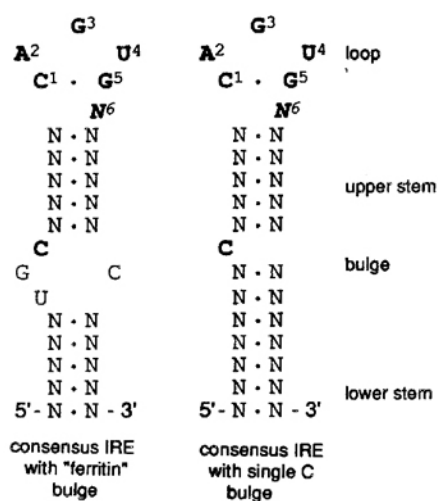


Figure I.10 : Motif IRE dans les ARNm. Une analyse comparative de différents IRE a permis de révéler la conservation d'une séquence consensus et d'une structure en épingle à cheveux. D'après Hentze et Kühn (1996).

Les IRP reconnaissent donc des séquences de nucléotides dans les régions non traduites, les UTR, de certains ARNm. Ces motifs IRE adoptent tous une structure commune en tige boucle et bien qu'une partie de leur séquence soit conservée (CAGUGU/C) il en existe de différents types (Figure I.10). Ces variations sont le reflet d'autant de différentes régulations dépendantes du fer qui peuvent ainsi fournir une plus grande finesse de réglage dans les réponses (Theil, 2000 ; Ke et al., 1998).

La ferritine (Ft) présente, à la fois pour sa sous-unité L que pour sa sous-unité H, des motifs IRE dans les UTR en 5' de ses ARNm près du codon initiateur AUG. Ces IRE sont très efficacement liés par les IRP et sont, de ce fait, l'objet d'une forte régulation selon la disponibilité en fer de la cellule (Theil, 2000). Les IRP, une fois fixées sur l'ARNm de la Ft, empêchent les interactions entre la petite sous-unité ribosomale (Muckenthaler et al., 1998) et le facteur d'initiation 4F, l'eIF4F, comprenant la CBP (*Cap Binding Protein*) qui est responsable de la reconnaissance de la coiffe en 5' de l'ARNm, empêchant ainsi la formation du complexe nécessaire à la traduction (Figure I.11). Des expériences de mutagenèse, *in vitro*, ont démontré l'importance de chacun des nucléotides impliqués dans la structure en tige/boucle ce qui a récemment été corroboré par l'identification de mutations spontanées chez des patients présentant le syndrome héréditaire d'hyperferritinémie, aboutissant à une cataracte précoce (Allerson et al., 1999). Les mutations au niveau du motif IRE de la sous-unité L de la ferritine entraîne une forte augmentation des niveaux de Ft dans le sérum et ce malgré l'absence d'une surabondance en fer. Le rapport entre l'effet de la mutation sur la reconnaissance des IRP, le niveau de Ft circulant et la gravité clinique de la maladie a été établi bien que le lien moléculaire entre le niveau élevé en ferritine dans le sérum et l'apparition de la cataracte soit encore inconnu (Cazzola et al., 1997 ; Allerson et al., 1999).

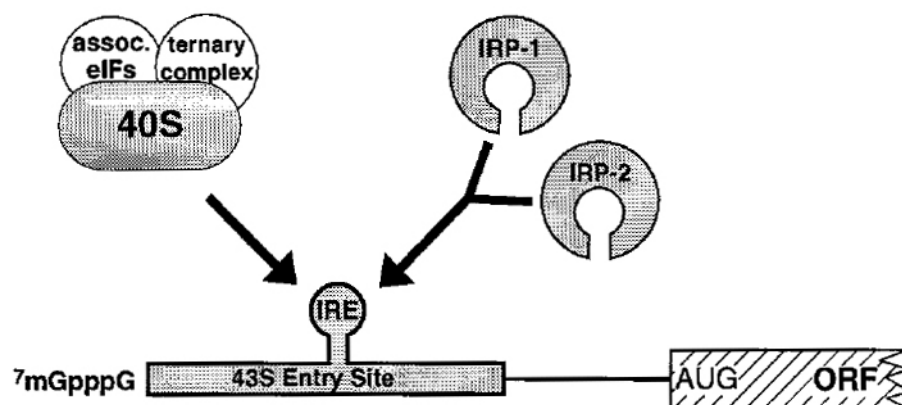


Figure I.11 : Régulation de la traduction par le système IRP/IRE. La fixation d'une des protéines IRP (1 ou 2) empêche la petite sous-unité ribosomale de venir se fixer sur l'ARNm ce qui permet de bloquer la traduction et d'inhiber l'expression du gène ainsi codé. D'après Hentze et Kühn (1996).

L'aconitase mitochondriale a aussi un IRE dans l'UTR en 5' de son ARNm ce qui lui permet d'être régulée par les variations intracellulaires en fer. Sa régulation est cependant plus limitée que celle de la Ft (Schalinske et al., 1998) ce qui est sûrement le fruit d'une différence des structures dans l'IRE (Ke et al., 1998). Comme le citrate est en mesure de lier le fer, les changements de concentration en citrate, dûs à l'action de l'aconitase qui catalyse la transformation du citrate en isocitrate, peut moduler la disponibilité du métal et ainsi affecter l'homéostasie du fer. De plus, les variations du niveau d'expression de l'aconitase mitochondriale, régulées par les IRP, peuvent présenter un moyen de coupler l'homéostasie du fer à la production d'énergie cellulaire comme cela a pu être suggéré par des études chez les procaryotes (Tang et al., 1999). L'existence d'un lien entre le cycle de Krebs et le métabolisme du fer a aussi été montrée par la découverte du fait que le transcrit de la succinate déshydrogénase b, qui oxyde le succinate en fumarate au cours du cycle de l'acide citrique chez la *Drosophila melanogaster*, contient aussi un motif IRE dans son extrémité 5'. Cela confère donc à cette enzyme une régulation dépendante du niveau cellulaire en fer selon le principe du système IRP/IRE (Kohler et al., 1995 ; Gray et al., 1996).

Dans un même genre, la synthèse de l'erythroïde aminolevulinate synthase, eALAS, une enzyme clé dans la biosynthèse de l'hème, est couplée à la disponibilité en fer dans la cellule par des interactions IRP-IRE (Cox et al., 1991 ; Dandekar et al., 1991) bien que pour cette protéine le niveau de contrôle de la traduction soit plus modéré (Mikulits et al., 1999). Une séquence conservée, typique des motifs IRE, est aussi présente dans les UTR en 5' des ARNm codant pour la ferroportine, un transporteur impliqué dans l'exportation du fer depuis les enterocytes et les macrophages (Donovan et al., 2000).

Bien que tous les ARNm précédents aient un IRE dans leur région 5', l'ARNm du récepteur à la transferrine, qui régule l'incorporation du fer par internalisation de l'holotransferrine, a cinq motifs IRE dans son extrémité 3' non traduite (Hentze et Kühn, 1996). Les interactions des IRP avec un nombre aussi important d'IRE n'affectent en rien la traduction de l'ARNm et au contraire cela augmente son temps de demi vie en le protégeant des activités RNases (Figure I.12). En revanche, l'augmentation de la stabilité des transcrits aboutit à une expression plus importante du TfR au niveau de la membrane cytoplasmique de la cellule et donc à une plus grande incorporation du fer. D'un autre côté, aucun motif de type IRE n'a été trouvé dans les UTR du transcrit du TfR2 ce qui peut être interprété comme une absence de régulation de ce gène par le fer (Fleming et al., 2000 ; Kawabata et al., 2000).

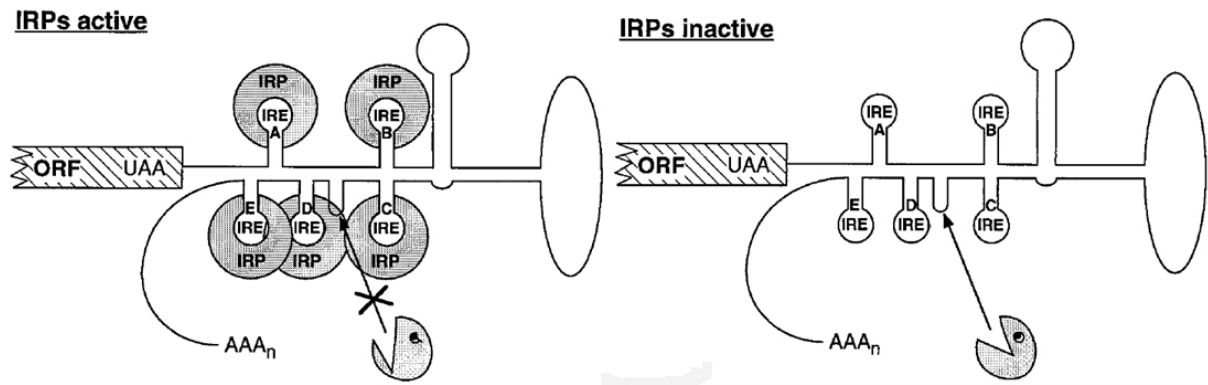


Figure I.12 : Régulation de la traduction par le système IRP/IRE. La fixation des IRP (1ou 2) empêche l'accessibilité des RNases à l'extrémité 3' des ARNm, augmentant leur temps de demi vie. D'après Hentze et Kühn (1996).

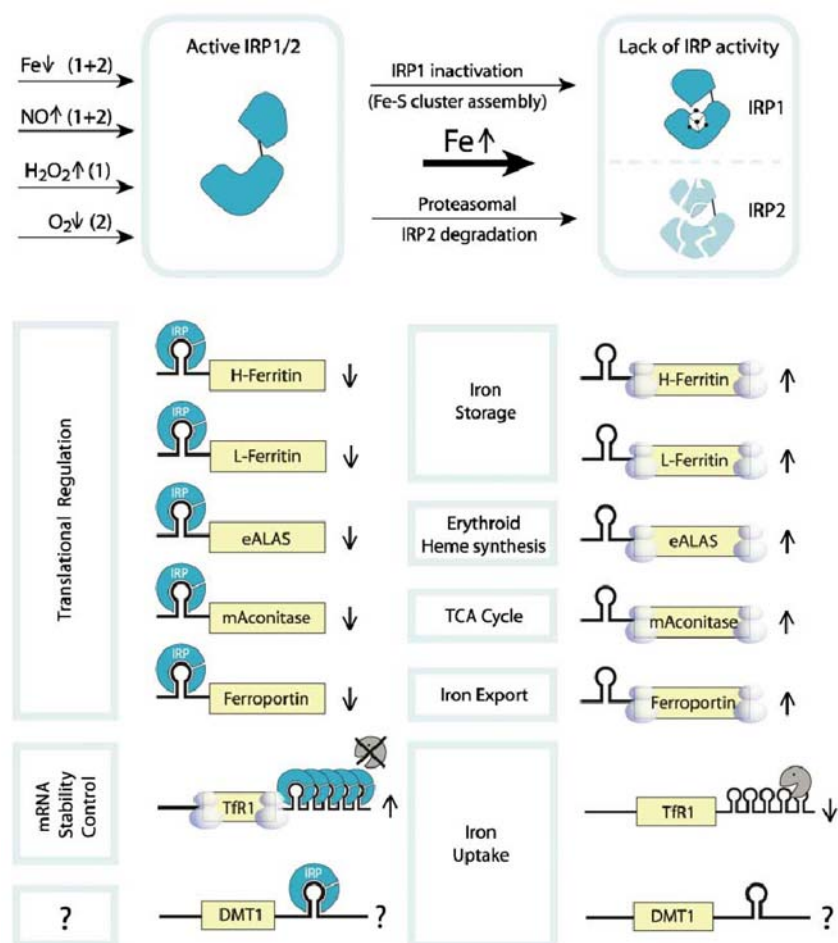


Figure I.13 : Système de régulation IRP/IRE. Les protéines impliquées dans le stockage du fer, la synthèse de l'hème, le cycle de l'acide citrique, l'exportation ou encore l'absorption du fer sont régulées, de façon coordonnée, par l'interaction des IRP avec les IRE. Les IRP sont régulées par différents stimuli mais les deux isoformes sont sensibles à la concentration en fer disponible dans la cellule. IRP1 est néanmoins plus sensible au stress oxydant alors que IRP2 est plus sensible au manque d'oxygène. D'après Hentze et al. (2004).

Un IRE potentiel est aussi présent dans l'UTR en 3', selon un épissage alternatif, de l'ARNm de l'un de deux transcrits de DMT-1, qui est le principal transporteur de fer dans les

cellules eucaryotes et plus spécialement dans les enterocytes (Gunshin et al., 1997). Cependant, cet IRE est probablement un compétiteur bien moins efficace pour IRP que celui de la transferrine (Wardrop et Richardson, 1999). En effet, la protéine codée par cet isoforme portant un IRE a été montrée comme étant modulée par le fer alimentaire dans le duodénum (Canonne-Hergaux et al., 1999) mais aucun effet n'a été trouvé sur les niveaux de DMT-1 dans des lignées cellulaires de fibroblastes (Wardrop et al., 1999). En bref, il y a une absence frappante d'évidence expérimentale directe sur le fait que cet IRE confère un contrôle quelconque du fer sur DMT-1 (Figure I.13).

Des expériences *in vitro* de SELEX (*Selective Enrichment of Ligands by EXponential amplification*) ont permis l'identification d'IRE, à partir de sondes dégénérées, capables de lier les IRP (Henderson et al., 1996 ; Butt et al., 1996 ; Henderson et al., 1994). Quelques-unes ont montré des interactions préférentielles soit avec IRP1 soit avec IRP2 (Menotti et al., 1998) mais aucun transcrit naturel portant ces IRE alternatifs n'a pu être trouvé à ce jour. De plus, il a été montré que dans le cas de la glycolate oxydase la présence d'un motif conservé IRE n'était pas toujours suffisant à la régulation par les IRP (Kohler et al., 1999).

2.3 Régulation des IRPs

Le simple modèle d'IRP comme étant des protéines bifonctionnelles avec des fonctions métaboliques distinctes dépendant seulement de la disponibilité en fer pour assembler un agrégat pourrait être correct pour la simple régulation d'IRP1 mais sûrement pas pour IRP2 qui n'a pas d'agrégat fer soufre et qui pourtant joue un rôle biologique très important. De plus, des résultats ont montré l'existence d'espèces moléculaires d'IRP1 incapables de fixer les motifs IRE, ni même d'avoir une fonction aconitase à cause de leur forme inactive soit de par leur agrégat [3Fe-4S] soit de par l'oxydation des résidus cystéines. Les dernières découvertes ont principalement été obtenues par l'étude des effets de différents agents et conditions pathologiques sur l'activité des IRP et il en résulte que le fer n'est pas la seule clé de régulation du système IRP/IRE.

2.3.1 Le stress oxydant

Les espèces réactives de l'oxygène, ou ROS pour *Reactive Oxygen Species*, qui sont inévitablement et constamment générées dans toutes les cellules durant leur vie normale sous des conditions aérobies, peuvent causer des dommages considérables à un grand nombre de molécules cibles essentielles à l'homéostasie cellulaire. Cependant, à des concentrations contrôlées, elles jouent aussi un rôle dans certaines voies physiologiques, comme la

transduction du signal, la signalisation cellulaire, la régulation des activités d'oxydoréduction, la prolifération cellulaire ou encore l'apoptose. La variation dans la disponibilité en fer est le principal moyen par lequel les cellules conservent un contrôle strict sur les niveaux de ROS. En effet, la séquestration appropriée du fer peut permettre un rôle physiologique, relativement sans danger, à l'anion superoxyde (O_2^-) et au peroxyde d'hydrogène (H_2O_2) sans prendre part pour autant à la production du très réactif radical hydroxyle (OH^\bullet) selon la réaction de Fenton (voir § I.1.1). Il n'est donc pas une surprise que l'activité IRP, en temps que maître régulateur du métabolisme cellulaire du fer, soit altérée en cas de stress oxydant. La Ft et le TfR ayant pour rôle le stockage et l'importation du fer dans la cellule, il paraît clair que leur régulation respective, dans ce genre de condition, doit être propice à limiter la disponibilité fer en contrôlant l'expression de ces deux protéines.

En effet, il a été démontré dans de nombreuses lignées cellulaires que, selon les conditions stressantes auxquelles elles étaient soumises, la synthèse de la ferritine était considérablement augmentée (Balla et al., 1992 ; Vile et al., 1993 ; Kwak et al., 1995 ; Lobreux et al., 1995). La réponse au stress oxydant s'en est donc retrouvée d'autant réduite dans ces cellules du simple fait de la surexpression de la sous-unité H de la Ft (Epsztejn et al., 1999). Cependant cet effet n'est pas cohérent avec l'activité IRP dans le cas de cellules traitées avec du peroxyde d'hydrogène, car le pool de fer labile augmente au lieu de diminuer ce qui risque d'exacerber la toxicité des ROS (Martins et al., 1995 ; Pantopoulos et al., 1995). Une analyse minutieuse des mécanismes et des signaux sous-jacents à l'activation des IRP dans leur rôle de fixation à l'ARN, sous ces conditions, a mis en évidence que cet effet rapide se produisait en réponse à une voie de signalisation de type cascade de phosphorylation, pouvant provenir des effets du peroxyde d'hydrogène sur la stimulation de la croissance cellulaire que du stress oxydant lui même (Burdon et al., 1996, Mueller et al., 2001). En effet, ni l'augmentation intracellulaire du niveau de H_2O_2 (Pantopoulos et al., 1997), ni le traitement d'extrait cytosolique avec du H_2O_2 ne conduisent à aucun effet stimulant sur l'activation des IRP (Martins et al., 1995 ; Pantopoulos et al., 1995 ; Cairo et al., 1996 ; Bouton et al., 1996). Afin de restaurer une réponse de la part des IRP, il faut impérativement ajouter au milieu expérimental des fragments de membrane, de l'ATP et du GTP (Pantopoulos et al., 1998). Il semble évident maintenant que des événements de phosphorylation sont impliqués dans l'activation des IRP et plus particulièrement d'IRP1. De plus, cette cascade de transduction du signal n'est possible que par l'ajout d'extrait membranaire, ce qui implique qu'un des éléments essentiels aux premières étapes de phosphorylation se trouve associé à la membrane cytoplasmique. A ce jour, il semble admis que la protéine kinase C puisse être impliquée dans

ces événements de phosphorylation mais son rôle exact n'est pas encore clair, de même que l'enchaînement des acteurs impliqués dans cette voie de signalisation n'est toujours pas bien défini (Pantopoulos et al., 1998, Mueller et Pantopoulos, 2002). Afin d'appuyer cette idée, il a été montré que la génération simultanée de peroxyde d'hydrogène et d'anion superoxyde, ce qui est le plus vraisemblable de se produire sous des conditions de stress oxydant que la production de H_2O_2 seul, inhibe l'activité IRP (Cairo et al., 1996). L'inactivation des IRP a aussi été corroborée par le traitement des cellules à la menadione (Gehring et al., 1999), une quinone cyclique qui produit du peroxyde d'hydrogène et des anions superoxydes, et il a été montré que, *in vivo*, la production de ROS, obtenue par traitement chimique (Smith et al., 1998 ; Cairo et al., 1995), inhibait aussi l'activité IRP.

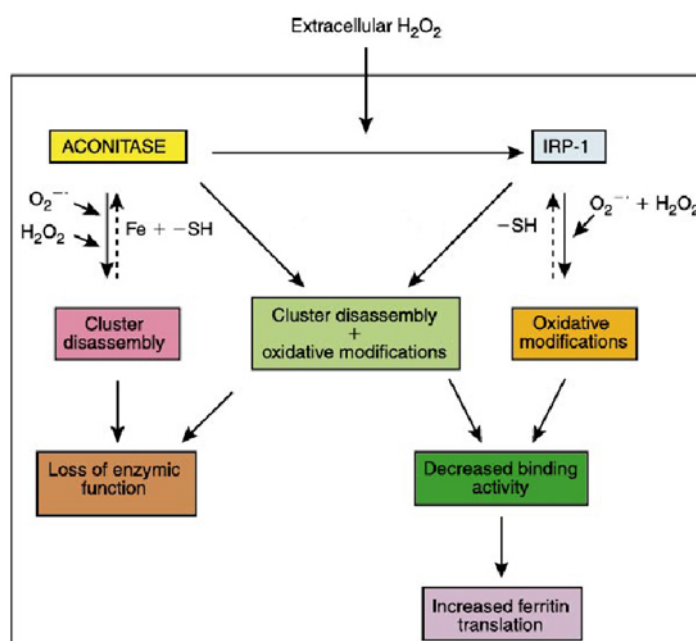


Figure I.14 : Régulation redox d'IRP1 et de l'aconitase mitochondriale. Quand l'anion superoxyde et le peroxyde d'hydrogène sont produits séparément dans la cellule l'aconitase perd, de façon réversible, sa fonction aconitase sans pour autant lui conférer une activité IRP1. En revanche quand les deux sont présents ensemble alors l'activité IRP1 est régulée de façon négative mais sans gêner l'activité aconitase. D'un autre côté l'ajout de H_2O_2 extracellulaire amorce une cascade de signalisation qui favorise l'activité de fixation d'IRP1 sur l'ARN au détriment de son activité enzymatique d'isomérase. D'après Cairo et Pietrangelo (2000).

Toutes ces considérations se réfèrent principalement à IRP1 qui est doté d'un agrégat $[4Fe-4S]$ alors qu'il y a peu de données concernant les effets des ROS sur IRP2. Le mécanisme de régulation d'IRP2, c'est-à-dire via une dégradation ubiquitine dépendante de la protéine par le protéasome après oxydation des résidus du domaine de dégradation (Iwai et al., 1998), suggère que la protéine IRP2 est hautement sensible aux changements redox de la cellule. En effet, IRP2, qui n'est pas activée par du peroxyde d'hydrogène extracellulaire (Menotti et al., 1998 ; Pantopoulos et al., 1996) est inhibée dans le foie de rat après une

exposition au stress oxydant (Cairo et al., 1995 et 1998 ; Tacchini et al., 1997 ; Mueller et al., 2001) et dans les cellules traitées à la menadione (Gehring et al., 1999).

En bref, des travaux récents ont offert de nouvelles perspectives concernant le rôle des IRP dans le stress oxydant ce qui a permis de réconcilier sa fonction dans une structure cohérente destinée à limiter la disponibilité du fer catalytiquement actif (Figure I.14)

2.3.2 Le monoxyde d'azote (NO)

Le fait que les agrégats fer soufre soient des cibles préférentielles du monoxyde d'azote et que l'inhibition de l'aconitase mitochondriale soit impliquée dans l'effet cytotoxique de NO (Drapier et Bouton, 1996) a provoqué un grand nombre de recherches pour vérifier si cette molécule, qui est connue pour être impliquée dans une grande variété de processus physiologiques et pathologiques, jouait ou non un rôle dans la régulation du métabolisme du fer. Il a ainsi rapidement été établi un lien direct entre la variation du niveau intracellulaire en fer et l'induction de la transcription du gène de la nitrique oxyde réductase (Weiss et al., 1994). Il a été aussi clairement démontré que NO avait un grand nombre d'effets sur IRP1 et IRP2 quelle que soit l'espèce oxydée de monoxyde d'azote utilisé. La variation de l'activité IRP, induite par NO dans les macrophages, est probablement la plus pertinente des conditions physiopathologiques, car c'est ce qui est généralement retrouvé dans les inflammations ou encore les anémies chroniques, et c'est ce qui a permis de caractériser de façon très complète et pertinente les interactions IRP-NO (Drapier, 1997).

En accord avec les premiers travaux réalisés sur l'augmentation de l'activité de liaison à l'ARN des IRP faisant suite à une stimulation par NO (Drapier et al., 1993 ; Weiss et al., 1993), des résultats plus récents ont mis en évidence que l'activité d'IRP1 était répétitivement retrouvée en constante augmentation après traitement de lignées de macrophage avec des cytokines, aboutissant à la production de monoxyde d'azote (Bouton et al., 1996), mais aussi avec des supports chimiques générant de NO (Mulero et Brock, 1999 ; Kim et Ponka, 1999) mais sans pour autant apporter aucune précision sur les mécanismes mis en jeu. Il n'est toujours pas clair si l'effet activateur de NO dépend d'une interaction directe avec l'agrégat d'IRP1, qui pourrait impliquer son désassemblage en passant par une étape [3Fe-4S] jusqu'à l'acquisition de l'apoforme capable de lier l'ARN, ou un effet lent du NO sur le pool intracellulaire de fer libre (Bouton et al., 1997 ; Oliveira et al., 1999 ; Kennedy et al., 1997 ; Wardrop et al., 2000). La destruction de l'agrégat est probablement insuffisante parce que les espèces dérivées de NO peuvent causer la formation d'IRP oxydée et donc inactive ce qui

implique obligatoirement sa réduction par la thiorédoxine (Oliveira et al., 1999) ou tout autre agent réducteur, afin de lui permettre d'acquérir une activité complète de fixation à l'ARN.

A l'inverse, IRP2 ne se comporte pas du tout comme IRP1. En effet, les macrophages de souris traités avec des lipopolysaccharides (LPS) et des interférons γ (IFN γ) produisent du NO (Weiss et al., 1993) ce qui entraîne une inhibition complète d'IRP2 (Recalcati et al., 1998 ; Bouton et al., 1998) et une induction contrôlée de la synthèse de la ferritine par LPS/IFN γ (Recalcati et al., 1998). L'inactivation d'IRP2 a été confirmée par des études qui ont aussi décrit l'effet préférentiel sur IRP2 de la forme oxydé de NO (NO⁺) (Kim et Ponka, 1999) et un effet dose-dépendant de l'IFN γ (Mulero et Brock, 1999). La dégradation d'IRP2 est probablement le mécanisme d'inhibition qui est induit par NO dans les macrophages (Bouton et al., 1998 ; Kim et Ponka, 2000).

Un nombre important d'incertitudes subsiste concernant les effets de NO sur l'activité IRP. Ce qui est en particulier dû à la chimie intrinsèque de la molécule de monoxyde d'azote qui implique que sa réactivité dépende de son état de réduction et donc de l'environnement dans lequel il est généré. Ceci est illustré par le fait que IRP1 traité avec du peroxy-nitrite, un produit peu commun obtenu à partir de NO et de O₂⁻, perd son activité aconitase sans gagner son activité de fixation à l'ARN (Bouton et al., 1997). De plus, beaucoup d'informations relatives aux interactions IRP-NO ont été obtenues par traitement des extraits cellulaires avec des donneurs de NO. Bien que ce soit un moyen pratique d'explorer les mécanismes moléculaires sous-jacents aux effets de NO sur l'activité IRP, il n'en demeure pas moins que c'est difficilement reproductible *in vivo*. En effet, toutes les réponses, *in vitro*, au NO nécessitaient obligatoirement l'addition de thiorédoxine (Oliveira et al., 1999).

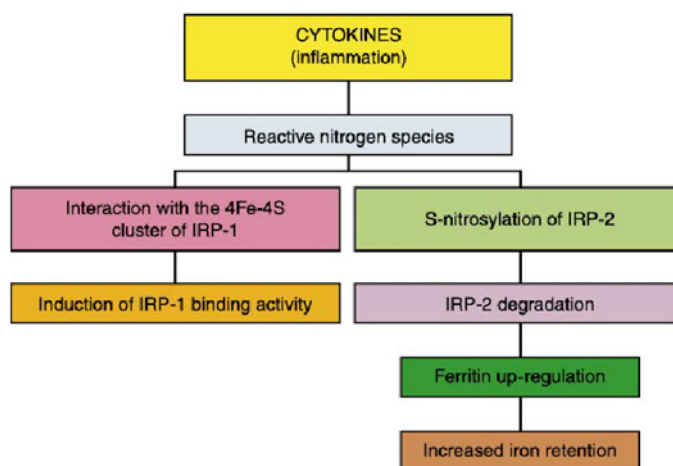


Figure I.15 : Modifications post-transcriptionnelles dans le métabolisme du fer des macrophages lors des phénomènes inflammatoires. L'activation induite par les cytokines module, à travers la voie NO, de façon différente les activités d'IRP1 et d'IRP2. D'après Cairo et Pietrangelo (2000).

Cependant, pour résumer les implications physiologiques de la stimulation de la voie NO sur l'activité IRP, il semble possible de conclure qu'en dépit de l'activation d'IRP1, la perte d'IRP2, qui est fortement exprimée dans les macrophages, peut-être pathologiquement plus importante car l'activation des macrophages par les cytokines est accompagnée par la stimulation de la synthèse de la ferritine (Recalcati et al., 1998 ; Kim et Ponka, 1999) et par une diminution du nombre d'ARNm codant pour le TfR (Pantopoulos et Hentze, 1995 ; Mulero et Brock, 1999 ; Kim et Ponka, 1999). La grande affinité d'IRP2 pour ses motifs cibles IRE peut au moins aider à expliquer cet effet dominant bien que IRP1 puisse aussi jouer un rôle via le monoxyde d'azote (Oliveira et Drapier, 2000).

Ces découvertes ouvrent de nouvelles perspectives dans la compréhension des mécanismes moléculaires impliqués dans la rétention du fer dans les macrophages qui est une des caractéristiques des problèmes inflammatoires (Figure I.15).

2.3.3 L'hypoxie

La découverte du fait que la Tf (Rolfs et al., 1997) et le TfR (Toth et al., 1999 ; Tacchini et al., 1999) soient codés par deux gènes inductibles par l'hypoxie a renforcé l'hypothèse d'un lien étroit entre le métabolisme du fer et celui de l'oxygène. Les IRP sont elles-mêmes sujetes à une régulation par l'oxygène mais des résultats contradictoires concernant les effets de l'hypoxie sur l'activité IRP ont été rapportés : deux études ont montré la diminution de l'activité IRP durant l'hypoxie (Hanson et Leibold, 1998 ; Tacchini et al., 1999) alors qu'une autre a montré une nette augmentation (Toth et al., 1999). D'une manière intéressante, le fait que les IRP soient facilement séparables au moyen d'une analyse sur gel retard a permis à Leibold et collaborateurs de démontrer que l'hypoxie régula de manière différentielle l'activité de fixation des IRP aboutissant à une diminution de l'activité de fixation d'IRP1 au profit de l'augmentation de son activité aconitase et d'une augmentation de la fixation d'IRP2 (Hanson et al., 1999). Cette augmentation de l'activité d'IRP2 provient d'un mécanisme encore assez confus impliquant la stabilisation de la protéine grâce au facteur HIF-1 (*Hypoxia-Inducible Factor 1*) inductible par l'hypoxie (Semenza, 2000).

Les conséquences fonctionnelles des différents comportements des deux IRP sont encore obscures, mais, si cette modulation opposée, observée dans les cellules de souris, est maintenue dans les cellules humaines, c'est-à-dire une augmentation de l'expression du TfR et une réduction de l'accumulation de la Ft trouvée dans les cellules manquant d'oxygène (Toth et al., 1999), suggère qu'IRP2 jouerait un rôle prédominant par rapport à IRP1. D'un autre côté, des études ont indiqué que l'augmentation de TfR, durant l'hypoxie, est

transcriptionnellement induit par HIF-1 (Lok et Ponka, 1999 ; Tacchini et al., 1999) malgré la régulation négative des IRP (Tacchini et al., 1999). De plus, l'inactivation de l'activité IRP, induite par l'hypoxie, pourrait très bien expliquer l'augmentation de la synthèse de Ft trouvée dans les oligodendrocytes en hypoxie (Qi et al., 1995). Le sens réel de cette modulation différentielle reste encore à établir mais cela pourrait être connecté non seulement à la régulation des protéines connues comme jouant un rôle dans le métabolisme du fer mais aussi à d'autres ARNm spécifiques à l'adaptation à l'hypoxie et présentant des motifs IRE. Néanmoins, cet exemple s'ajoute à la liste des conditions et agents qui modulent spécifiquement les deux IRP dans des directions opposées.

2.3.4 La croissance cellulaire

La synthèse *de novo* des protéines contenant du fer, nécessaires à la prolifération cellulaire, comme par exemple la ribonucléotide réductase qui fournit les désoxyribonucléotides pour la synthèse de l'ADN, peut limiter la disponibilité du fer depuis son pool cellulaire et ainsi avoir une action d'activation des IRP. Ainsi un haut niveau d'activité de fixation des IRP a été détecté dans des cellules mitotiques (Testa et al., 1991 ; Teixeira et Kühn, 1991 ; Seiser et al., 1993) et, *in vivo*, après une hépatectomie partielle (Cairo et Pietrangelo, 1994). Néanmoins, il est à remarquer que la haute activité de fixation d'IRP2, durant la régénération du foie, est accompagnée par une augmentation des niveaux d'ARNm du TfR mais pas d'une réduction de l'expression de la Ft. Il a été montré que l'augmentation de la synthèse de la Ft, sous ces conditions, est due non seulement à l'activation de la transcription du gène de la Ft mais aussi à la diminution de l'efficacité de l'activité répressive des IRP alors que leur rôle de stabilisateurs des ARNm demeure non affecté (Cairo et al., 1998). Une explication possible pour ce phénomène est la grande disponibilité du complexe d'initiation eIF4F qui entre en compétition active avec les IRP durant la régénération du foie (Aloni et al., 1992). Cette opposition du paradigme qui coordonne et oppose le contrôle de l'expression de la Ft et du TfR montre une fois de plus que la machinerie IRP/IRE n'est qu'une partie d'un système de régulation très complexe et qui interagit avec d'autres composants de la cellule qui ont eux aussi un rôle de régulateur. Cette hypothèse a été supportée par les résultats d'une étude démontrant que les oncoprotéines régulant le passage de la croissance à la différenciation de certaines cellules, affectaient aussi l'expression de la Ft et de l'eALAS régulée par les IRP.

De nouvelles découvertes ont montré que l'induction d'IRP2 dans les cellules en croissance n'était pas seulement le fruit d'une diminution du niveau de fer mais celui d'une

activation transcriptionnelle spécifique due à la surexpression de l'oncogène c-myc (Wu et al., 1999). Ce résultat semble renforcer l'idée qu'une multiplication cellulaire efficace requière un pool de fer important. IRP2 pourrait être, de plus, la cible des oncogènes afin de répondre aux exigences spécifiques du métabolisme d'une cellule cancéreuse en leur donnant ainsi un avantage sélectif non négligeable sur le tissu sain (Dang et Semenza, 1999).

3 Objectifs du travail de thèse

Il est maintenant clair que la régulation du fer est quelque chose de primordiale à tout point de vue et qu'une compréhension aussi précise que possible des phénomènes mis en jeu ne peut être qu'un avantage dans l'anticipation de problèmes aussi importants que les maladies neurodégénératives ou la croissance et la prolifération des cellules tumorales.

Le système sur lequel nous avons décidé de nous pencher semble être le point de passage obligé de toutes les protéines impliquées de près ou de loin dans l'homéostasie du fer chez les vertébrés. En effet, il semble clair que les protéines IRP1 et IRP2 sont les deux seuls maîtres régulateurs de tout un ensemble aussi complexe et fin que celui de l'incorporation, du stockage, de l'exportation ou encore de l'utilisation du fer au sein des cellules, des tissus et d'une façon beaucoup plus large de l'organisme. Bien que ces deux protéines paraissent pouvoir marcher de paire, il est maintenant établi que IRP2 est le pilier central de la régulation de l'homéostasie du fer alors que IRP1 se voit plutôt dévolu des rôles plus précis mais néanmoins primordiaux dans le développement cellulaire comme ce peut être le cas de la réponse au stress oxydant.

Dans le but de démarrer cette étude nous avons décidé de nous lancer dans la course en travaillant sur la protéine IRP1 avec comme intention de résoudre sa structure tridimensionnelle par diffraction aux rayons X. Ce choix s'est vu orienté par le fait qu'IRP1 est exprimée dans tous les tissus et dans toutes les cellules de l'organisme à des concentrations bien supérieures à celles d'IRP2 mais aussi par le fait que nous pouvions nous appuyer sur l'expertise de nos collaborateurs du Département Réponse et Dynamique Cellulaire du Commissariat à l'Energie Atomique de Grenoble, qui travaillent sur ce système et qui ont pu acquérir toutes les compétences nécessaires aux différentes techniques de purification et de caractérisation de la protéine.

Dans un premier temps, nous avons donc concentré nos efforts sur l'obtention de cristaux exploitables. Les difficultés ont commencé dès le début du projet, où le simple fait de

travailler sur IRP1 sous sa forme aconitase, nous imposait l'utilisation d'une atmosphère anaérobie et donc l'emploi d'une boîte à gants.

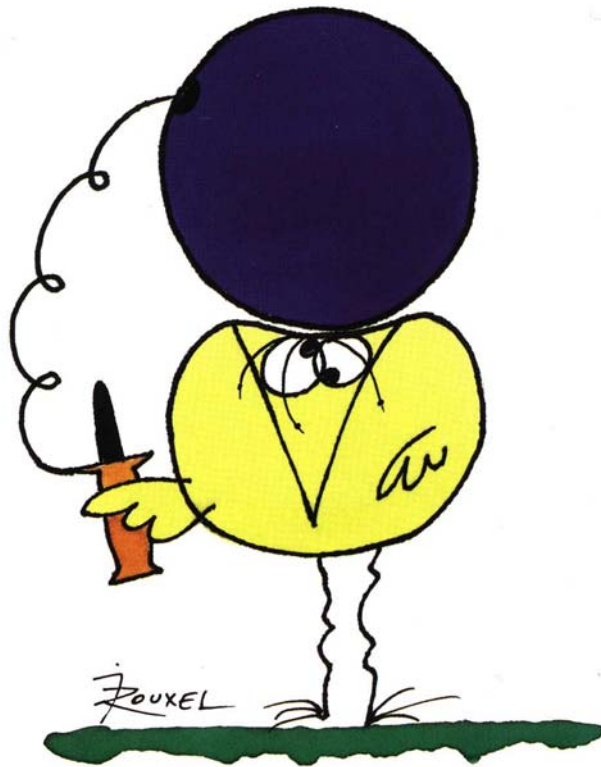
La seconde étape a été la plus longue et surtout la plus délicate, à savoir la résolution du problème de la phase. IRP1 est une protéine de gros poids moléculaire et la présence de l'agrégat [4Fe-4S] n'a pas été suffisante pour résoudre la structure de la protéine par les méthodes SAD ou MAD. De même, malgré l'homologie avec l'aconitase mitochondriale de bœuf nous n'avons pas pu utiliser la technique du remplacement moléculaire et nous avons dû nous orienter vers la recherche de dérivés d'atomes lourds. Une fois les phases obtenues nous avons commencé la construction manuelle du premier modèle jusqu'à le faire évoluer, via plusieurs cycles d'affinement, vers la structure finale que nous connaissons maintenant.

La troisième étape a consisté dans une étude de comparaison avec l'aconitase mitochondriale de bœuf afin de dégager les grandes différences pouvant peut-être expliquer pourquoi l'aconitase mitochondriale n'a pas cette capacité à fixer l'ARN mais surtout mettant en exergue toutes les subtilités de la fonction aconitase chez IRP1.

Enfin la dernière partie de cette étude a consisté, sur la base des résultats cristallographiques, à essayer de dégager quelques idées sur ce que peut être le mode de fixation du motif IRE sur la protéine IRP1 de même qu'à essayer d'anticiper les variations structurales qui pourraient être retrouvées sur la protéine IRP2.

Matériel et Méthodes

Les devises Shadok



EN ESSAYANT CONTINUUELLEMENT
ON FINIT PAR RÉUSSIR. DONC:
PLUS ÇA RATE, PLUS ON A
DE CHANCES QUE ÇA MARCHE.

Chapitre II : Matériel et méthodes

La résolution d'une structure tridimensionnelle par cristallographie nécessite trois étapes indispensables. La première consiste en l'obtention d'un échantillon aux qualités d'homogénéité et de pureté proches de la « perfection ». Je ne m'attarderai pas sur cette étape, mais je tiens à souligner le travail exceptionnel de nos collaborateurs Jean-Marc Moulis et de son ancien étudiant Xavier Brazzolotto. Ils ont su nous fournir des préparations d'une grande qualité pour nos essais de cristallisation. La deuxième étape est à proprement parlé l'obtention d'un monocristal exploitable et enfin la dernière consiste en la résolution du problème des phases. Pour commencer je vais aborder les principes de la cristallogénèse avant de passer aux complications qui se présentent lors de la résolution d'une structure cristallographique et des moyens qui existent pour surmonter ces problèmes.

1 La cristallogénèse

L'ensemble du travail cristallographique pour déterminer la structure d'une nouvelle protéine est entièrement conditionné par l'obtention de cristaux de taille et de qualité de diffraction suffisantes. La mise au point de nouvelles conditions de cristallisation et l'amélioration de celles-ci sont, aujourd'hui encore, menées de manière très empirique et sous la forme d'essais-erreurs répétés. Pour aider dans ce travail, un arsenal complet de recettes existe.

1.1 Les principes

Un cristal est l'arrangement périodique et ordonné d'un motif composé d'un ensemble de molécules de protéines se répétant à l'identique d'une maille à l'autre. Le principe de la cristallisation est donc d'induire un changement d'état de la protéine de sa phase soluble à une phase solide et ordonnée, ces deux états étant en équilibre. Ce phénomène de transition de phase dépend de nombreux paramètres et peut se diviser en deux étapes : la germination ou plus communément appelé nucléation et la croissance cristalline.

La difficulté consiste à trouver le subtil compromis entre les différents facteurs cinétiques et thermodynamiques qui contrôlent la solubilité de la protéine qui est tributaire des interactions entre les acides aminés de surface et le solvant. Pour induire la nucléation, il faut amener la protéine dans un état soluble instable de sursaturation en diminuant lentement sa solubilité (voir Figure II.1). Ce qui revient à déstabiliser les interactions protéines-solvant au profit d'interactions protéines-protéines. Des interactions non-spécifiques conduisent à la

précipitation (phase solide amorphe), au contraire des interactions spécifiques entre les molécules de protéines mènent à la formation de germes cristallins. Le retour à l'équilibre thermodynamique s'effectue alors plus ou moins rapidement par la croissance de la phase solide (les cristaux dans le meilleur des cas).

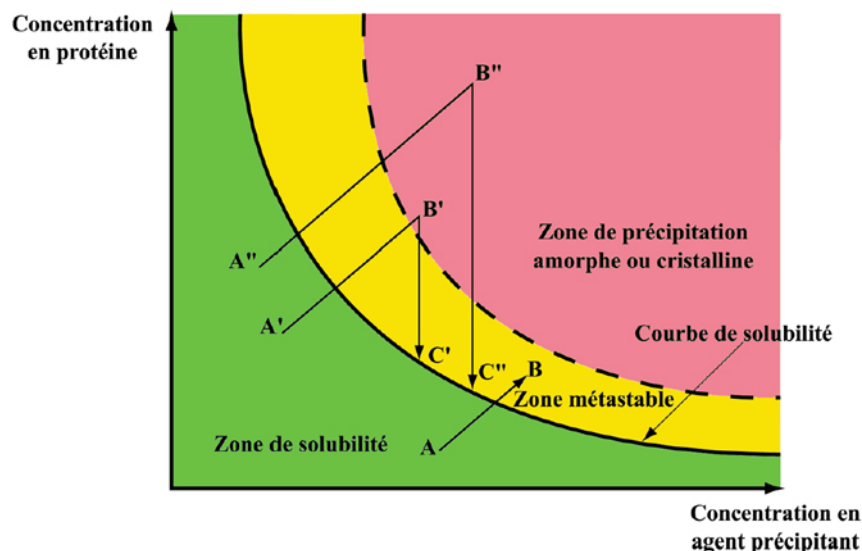


Figure II.1 : Différents chemins expérimentaux sur le diagramme de phase. A : la goutte va se concentrer, mais pas suffisamment pour permettre l'apparition de germes. La solution restera saturée à cause d'une cinétique trop lente. A' : la goutte va se concentrer jusqu'à atteindre une sursaturation suffisante pour permettre l'apparition de germes (B') qui sont dans le meilleur des cas cristallins. Les germes vont alors croître et la concentration en protéine va diminuer jusqu'à atteindre la limite de solubilité (C'). A'' : la goutte va se concentrer de façon à atteindre la zone de précipitation, loin de la limite de la zone métastable. L'énergie du système sera telle que de nombreux germes vont apparaître qui conduiront à une multitude de petits cristaux. Il est aussi possible que la concentration de la goutte soit telle qu'elle favorise, d'un point de vue cinétique, l'apparition d'un précipité amorphe.

1.1.1 Les paramètres physico-chimiques

Les paramètres qui influencent la solubilité des protéines sont très nombreux et variables d'une protéine à l'autre. En pratique, pour trouver un état de sursaturation qui conduise à la cristallisation, il faut tester et maîtriser plusieurs facteurs :

- La pureté et la concentration de la protéine : absence de contaminant, de micro-hétérogénéité ou d'isomérisation (homogénéité conformationnelle) ;
- Les agents agissant sur la solubilité, la stabilité et/ou la conformation de la protéine : tampons, ions, substrats ou inhibiteurs, cofacteurs, détergents mais aussi le temps écoulé depuis la purification (pendant lequel des dégradations peuvent avoir lieu) ;
- Le pH dont dépend la répartition des charges à la surface de la protéine ;
- La nature et la concentration des agents précipitants et des additifs ;
- La température affectant la solubilité, la stabilité de la protéine et la cinétique de cristallisation.

1.1.2 Les principaux agents précipitants

L'agent précipitant tient un rôle majeur dans le processus de cristallisation puisqu'il affecte directement et de façon majoritaire la solubilité de la protéine. Les différents agents les plus couramment employés peuvent être classés en trois familles distinctes suivant leurs effets sur les interactions protéines solvant et protéines protéines :

- Les polymères qui agissent par exclusion du solvant au voisinage de la protéine (le PEG est le plus utilisé),
- les sels non chaotropiques (non dénaturants pour les protéines) qui modifient la force ionique et augmentent les interactions hydrophobes en excluant le solvant (le sulfate d'ammonium, les acétates ou les chlorures),
- les solvants organiques qui diminuent la constante diélectrique du milieu (MPD, isopropanol, dioxane...).

Afin de maîtriser pleinement la solubilité de la protéine, il est souvent nécessaire de combiner les agents précipitants ou d'y associer un ou plusieurs additifs comme les détergents ou les alcools.

1.2 Les techniques

Il existe de nombreuses méthodes de cristallisation des protéines, chacune permettant d'atteindre et de traverser différemment les zones de germination et de croissance cristalline : par équilibre en phase vapeur, par dialyse ou par « *batch* » (Ducruix et Giege, 1992).

1.2.1 La diffusion en phase vapeur

La technique la plus utilisée actuellement est celle de la diffusion de vapeur en goutte suspendue (Wlodawer et Hodgson, 1975 ; McPherson et al., 1995). Elle est basée sur les propriétés thermodynamiques qui imposent aux concentrations en agents précipitants de deux solutions, enfermées dans un milieu hermétiquement clos, de tendre vers un équilibre par diffusion de vapeur de solvant.

Elle est particulièrement adaptée aux essais réalisés sur de petits volumes (1 à 5 µl). En début d'expérience, une goutte de protéine est diluée avec une solution contenant un agent précipitant, au voisinage de la zone de sursaturation. Cette goutte est placée dans une enceinte hermétiquement scellée (par de la graisse minérale ou par un joint synthétique) en même temps qu'un large volume de la solution de précipitant, comme indiqué sur la Figure II.2. Le retour à l'équilibre des concentrations est atteint lentement par évaporation d'eau (diffusion de vapeur), de la goutte vers le réservoir. La diminution du volume de la goutte entraîne en

même temps l'augmentation des concentrations en précipitant et en protéines (voir Figure II.1). Le réservoir ne sert qu'à maintenir une pression de vapeur constante.

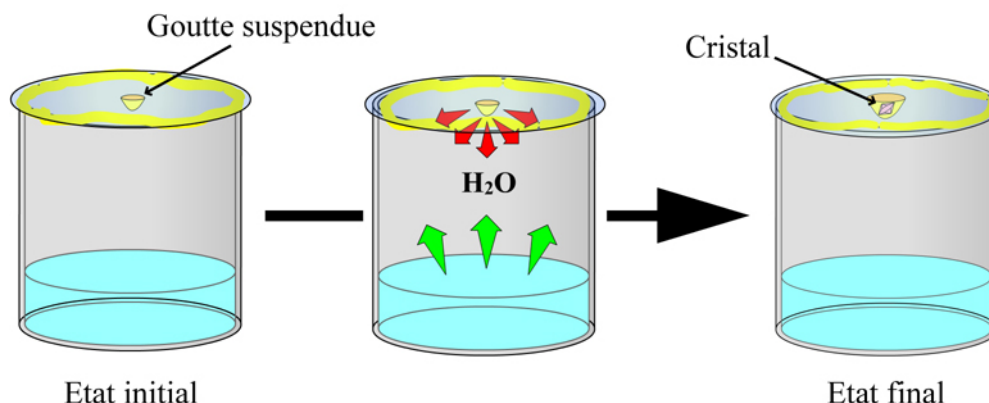


Figure II.2 : Technique de la goutte suspendue.

1.2.2 L'ensemencement

Lorsque les cristaux sont trop petits, il est recommandé d'utiliser la technique d'ensemencement. En effet, la germination demande plus de sursaturation que la croissance cristalline. Cette méthode consiste à utiliser les germes obtenus en forte sursaturation pour ensemer en faible sursaturation. Les germes peuvent être soit des monocristaux (macro-ensemencement) soit un broyat de cristaux (micro-ensemencement). En permettant de maîtriser le nombre de germes ensemencés, cette méthode permet de favoriser la croissance de quelques grands cristaux au lieu d'une multitude de petits. Par ailleurs, en dissociant les étapes de germination et de croissance, elle permet de faire pousser les cristaux dans des conditions plus variées et moins saturantes. Or les cristaux poussent souvent plus lentement et donc plus régulièrement en faible saturation.

1.3 La recherche des conditions de cristallisation

L'ensemble des paramètres physico-chimiques, qui peuvent influencer la cristallisation des macromolécules biologiques, définit un espace paramétrique de grande dimension. *A contrario*, la quantité d'échantillon dont nous disposons pour chercher et optimiser les conditions de cristallisation d'une nouvelle protéine est, quant à elle, souvent très limitée. En l'absence d'une base théorique ayant démontré une réelle efficacité pour orienter cette recherche, ce sont les méthodes qui permettent de tester le plus grand nombre de paramètres, tout en s'appuyant sur l'expérience acquise avec d'autres protéines, qui sont les plus utilisées. L'approche classique de ce travail est habituellement décomposée en trois étapes consécutives : le criblage, l'analyse du plan factoriel et l'optimisation.

1.3.1 Le criblage

En prenant comme base l'ensemble des conditions connues pour la cristallisation de protéines, plusieurs équipes ont proposé des matrices de criblage ou, en anglais, « screening » (Cudney et al., 1994 ; Jancarik et Kim, 1991 ; McPherson, 1992). Les deux matrices les plus utilisées sont celles de Jancarik et Kim (1991) et de Cudney et collaborateurs (1994), elles sont commercialisées par Hampton Research sous les noms de Crystal Screen I et II. Chacune contient une cinquantaine de conditions différentes. Le protocole de Jancarik et Kim contient des conditions « standards », alors que celui de Cudney et collaborateurs propose un cocktail de conditions utilisant des molécules un peu plus exotiques (Jeffamine, polyéthylèneimine, PEGMME...) ou peu utilisées en cristallisation lors des premières recherches (sels de métaux type césium, cobalt, fer...).

L'intérêt des criblages est de permettre de dégager rapidement, sans connaissance préalable, les premières tendances à la cristallisation du système en utilisant peu de matériel mais en testant un grand nombre de paramètres. Dans les cas favorables, elles permettent d'obtenir rapidement les premiers cristaux.

1.3.2 Les plans d'analyse factorielle

A partir des résultats obtenus par ces premiers essais, une analyse plus fine de l'effet de certains paramètres peut être orientée en élaborant des plans d'expériences qui seront soumis à une analyse statistique. Ces plans peuvent être complets, ce qui revient à tester toutes les combinaisons des paramètres fixés, ou incomplets, ce qui s'apparente à échantillonner de façon homogène l'espace des paramètres à explorer en un minimum d'expériences (Carter et Carter, 1979).

Le principe d'élaboration d'un plan factoriel incomplet est de combiner entre elles et aléatoirement les différentes valeurs (ou niveaux) de chaque variable (ou paramètre) de manière à ce qu'elles soient représentées statistiquement le même nombre de fois et d'éviter la redondance des expériences. Le nombre de conditions à tester dépend de l'espace multiparamétrique choisi et du nombre de valeurs attribuées à chaque variable. Si N représente le nombre de conditions d'un plan complet, il est possible de réduire jusqu'à \sqrt{N} le nombre minimum d'expériences pour un plan incomplet. Au laboratoire, le logiciel SAmBA (Audic et al., 1997) permet de déterminer ce genre de plan d'expérience.

En théorie, pour l'analyse statistique, un score (ou figure de mérite) est attribué par examen visuel à chaque résultat d'expérience et donc par conséquent à chaque niveau de chaque variable. Les scores ainsi attribués indiquent l'effet positif ou négatif de chaque

paramètre sur la cristallisation. Dans le cas le plus simple, l'échelle des scores possède trois niveaux : 0 pour les gouttes claires, 1 pour les précipités et 2 pour les cristaux.

Après avoir dégagé les effets majeurs, un nouveau plan peut-être élaboré en combinant les facteurs positifs. Ainsi par étapes successives, il est possible de réduire le nombre de conditions à explorer jusqu'à revenir à des plans complets afin d'affiner et d'améliorer les conditions de cristallisation.

1.3.3 L'optimisation

Souvent, les cristaux au terme d'essais systématiques de quelques paramètres ne permettent pas encore d'obtenir une diffraction de bonne qualité. Une étape d'optimisation est nécessaire. Le but de cette étape est d'ajuster finement les conditions afin de diminuer la germination, d'augmenter la taille des cristaux, d'améliorer leur aspect, d'éliminer les problèmes de macle ou de cristaux collés et surtout d'augmenter l'ordre cristallin. Différentes recettes peuvent être mises en œuvre pour limiter le taux de nucléation ou ralentir la croissance cristalline :

- le scale-up : dans de nombreux cas des cristaux plus gros peuvent être obtenus, moyennant quelques réajustements des conditions, en augmentant simplement la quantité des produits utilisés (volumes des gouttes) ;
- la modification du rapport volumique entre la solution de protéine et le réservoir permet de modifier la vitesse de concentration et d'atteindre un niveau de sursaturation différent ;
- la méthode de perturbation : cette approche consiste à modifier les conditions initiales de cristallisation afin de les déplacer vers des conditions plus favorables. Cela peut être réalisé en ajoutant des additifs chimiques de divers types. Ces additifs peuvent être des agents de solubilisation (les détergents non ioniques, le dioxane...), des sels chaotropiques (l'urée, le thiocyanate de potassium...), des sels métalliques (Zn^{2+} , Ni^{2+} ...), des composés agissant sur les potentiels d'oxydoréduction (DTT, ferricyanure...) ;
- l'ensemencement : cette méthode (voir § II.1.2.2) peut permettre de ralentir la croissance cristalline et d'éviter les contraintes de la germination ;
- la cristallisation en gel : la cristallisation dans les gels de silice ou d'agarose semble efficace pour, à la fois, réduire le taux de germination et augmenter la taille et la qualité de diffraction des cristaux (Robert et al., 1992). Les gels permettent d'éviter les effets de convection qui perturbent la formation des cristaux ;

- le contrôle de la diffusion de vapeur : différentes techniques permettent de modifier les échanges de vapeur entre la goutte et le réservoir, celle employée dans cette étude consiste à placer une couche d'huile à la surface du réservoir (Figure II.3). Cette dernière agit comme une barrière et les vapeurs d'eau diffusent plus lentement ce qui permet d'agir directement sur la cinétique de croissance des cristaux (Chayen, 1997). Cette technique permet de réduire le nombre de cristaux en croissance et d'augmenter leur taille.

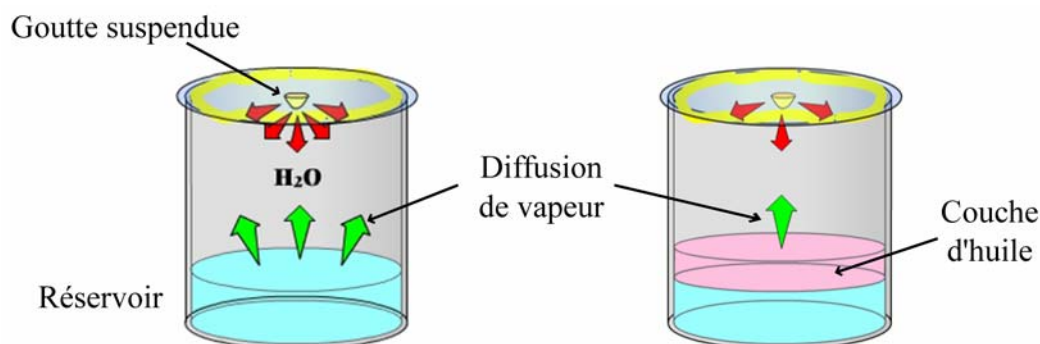


Figure II.3 : Limitation de la diffusion de vapeur grâce à l'emploi d'une barrière minérale constituée d'un mélange d'huiles plus ou moins perméable à l'eau. La diffusion de vapeur est moins importante après l'ajout de la barrière d'huile (schéma de droite) alors que sans elle est beaucoup plus importante (schéma de gauche).

1.4 La boîte à gants

La protéine sur laquelle j'ai travaillé a été produite dans le groupe de Jean-Marc Moulis, directement à partir de bactéries *E. coli* contenant le plasmide codant pour la forme humaine de la protéine IRP1. Toutes les étapes de purification ont été réalisées en absence d'oxygène du fait même de la sensibilité de l'agrégat [4Fe-4S] de la protéine. A cause de cette sensibilité prononcée pour l'oxygène et du fait que nous souhaitons travailler dans un premier temps sur la forme aconitase de la protéine, nous avons décidé de réaliser tous les essais de cristallisation dans une boîte à gants, sous atmosphère anaérobie.

1.4.1 Description

La boîte à gants du laboratoire comporte deux postes de travail pour la cristallisation et la manipulation des cristaux (Vernède et Fontecilla-Camps, 1999). Deux sas de taille différente, reliés à une pompe à vide, permettent de faire entrer ou sortir des objets sans contaminer l'enceinte par de l'oxygène. La boîte à gants a une atmosphère confinée, équilibrée avec de l'azote et un taux d'hydrogène variable autour des 10% durant les essais de cristallisation. La température est régulée par un groupe froid entre 20 et 22°C, de même que le taux d'humidité qui oscille aux alentours de 50%. Le taux d'oxygène est suivi par une sonde et l'utilisation d'un catalyseur au cuivre permet de maintenir ce taux à moins de 2 ppm.

La Figure II.4 représente la boîte à gants utilisée au laboratoire pour les essais de cristallisation des protéines sensibles à l'oxygène.



Figure II.4 : Vue de la boîte à gants. Le poste de gauche permet le stockage des boîtes de cristallisation alors que celui de droite est utilisé pour la manipulation des cristaux et la réalisation des gouttes.

1.4.2 Manipulation dans la boîte à gants

Afin d'éviter le relargage d'oxygène adsorbé sur le matériel de cristallisation, les boîtes de cristallisation Nextal[®] ainsi que les cônes des micropipettes sont laissés une nuit sous vide dans le sas avant utilisation. De même, pour éliminer l'oxygène dissous dans les solutions, celles-ci sont saturées par l'atmosphère de la boîte à gants par bullage pendant 20 min à l'aide d'une pompe à aquarium.

1.5 Cryocristallographie et congélation des cristaux

Lors de l'enregistrement de données de diffraction, les cristaux sont exposés plus ou moins longtemps aux rayons X. Ces rayonnements provoquent la dégradation des cristaux, principalement en générant des radicaux libres (radiolyse de l'eau) et des phénomènes d'ionisation (éjection des électrons de cœur par des ions métalliques au voisinage des seuils d'absorption). Avec l'emploi de source de rayonnement de plus en plus intense, le besoin de protéger les cristaux contre cette dégradation est devenu indispensable. La congélation à des températures inférieures à 140 K permet de diminuer la propagation des radicaux libres dans le cristal. Cela empêche également la congélation de l'eau sous forme cristalline laquelle provoque, par diffraction des rayons X, l'apparition de taches parasites sur les clichés de diffraction (principaux anneaux de glace à 3.5, 3.7 et 3.9 Å). Cette méthode permet en outre de faciliter la conservation et le transport des cristaux.

1.5.1 Le propane liquide

La congélation des cristaux se fait habituellement par trempage dans l'azote liquide. Toutefois, l'azote a une faible chaleur latente de vaporisation. Cela entraîne un phénomène de caléfaction lorsque le cristal est plongé dans l'azote, ce qui empêche une congélation rapide optimale surtout si le cristal est de grande taille. De plus, lors de l'entrée de l'azote dans la boîte à gants, celui-ci va s'évaporer pendant l'étape de vide à l'intérieur du sas. Pour résoudre ce problème, une technique de congélation rapide dans du propane a été utilisée (Vernède et Fontecilla-Camps, 1999). A -196°C , le propane trempé dans l'azote est solide. Cet état permet de le rentrer dans la boîte à gants avant fusion avec un minimum d'évaporation thermique. Lorsque l'état de fusion est atteint, le cristal est plongé dans le propane liquide qui se trouve loin de sa température d'ébullition. Il n'y a donc pas de phénomène de caléfaction et le cristal congelé peut être sorti de la boîte à gants avant que la température du propane n'atteigne -130°C qui correspond à sa température de vaporisation. Il n'y aura donc pas de transition glace vitreuse - glace cristalline. Le propane est alors re-solidifié dans l'azote liquide pour préparer un nouveau cycle de congélation en boîte à gants.

1.5.2 Le rôle du cryoprotectant

Les cristaux des macromolécules biologiques contiennent beaucoup d'eau (de 30 à 80%), ce qui peut entraîner la formation de glace cristalline lors de la congélation. Cette glace peut engendrer une augmentation du volume du cristal, ce qui le détériore, voire le détruit. Pour éviter cela, une méthode consiste à remplacer une partie de l'eau par un cryoprotectant selon une technique de trempage c'est-à-dire d'échange. Divers cryoprotectants peuvent être utilisés comme, par exemple le glycérol, le sucrose, le méthylpentanediol (MPD) ou le PEG 400. Par ailleurs, l'eau autour du cristal peut, elle aussi, cristalliser lors de la congélation, et ainsi perturber les clichés de diffraction par apparition de ce qui est appelé communément des anneaux de glace à 3,5, 3,7 et 3,9 Å. L'utilisation d'un cryoprotectant permet d'empêcher la formation de glace cristalline au profit d'un état vitreux amorphe et donc plus transparent aux rayons X.

2 La cristallographie

La cristallographie est un type de microscopie dans lequel les lentilles, permettant de produire l'image d'un objet à partir de la lumière qu'il diffuse, ont été remplacées par une transformation mathématique : la *transformée de Fourier* (TF), qui permet, elle, de calculer cette image à partir de la mesure directe de l'intensité diffusée.

2.1 La diffraction des rayons X par les cristaux

La diffraction met en évidence le caractère ondulatoire de la lumière. Cela s'explique par les interférences destructives ou constructives d'ondes cohérentes diffusées par les atomes d'un réseau lorsqu'ils sont exposés aux rayons X. L'arrangement périodique d'un cristal fait qu'un faisceau incident peut produire plusieurs faisceaux diffractés, appelé réflexions, ayant chacune une direction et une intensité déterminée. L'ensemble des **directions** de diffraction dépend uniquement de l'orientation des trois vecteurs de base **a**, **b** et **c** qui définissent le réseau cristallin, par rapport au faisceau incident (§ II.2.1.1). L'expérience de diffraction peut être décrite par les équations de Laue ou de Bragg. La **proportion** relative d'intensité diffractée dans chaque direction dépend (si les modulations induites par l'absorption ou la polarisation ne sont pas pris en compte), de la répartition de la densité de matière diffusante (dans le cas des rayons X, ce sont les électrons) à l'intérieur de la maille (§ II.2.1.2). La **symétrie** de l'arrangement cristallin impose des relations particulières sur l'intensité et le déphasage de certaines réflexions, qui sont appelées réflexions équivalentes. Elle peut aussi induire des extinctions d'intensité.

2.1.1 Les directions de diffraction

La diffraction par un réseau n'est observée que dans les directions de l'espace qui vérifient les **équations de Laue**. Si \mathbf{s}_0 est le vecteur d'onde incident et \mathbf{s} le vecteur d'onde diffusé (avec λ pour la longueur d'onde : $\mathbf{s}_0 = \mathbf{s} = 1/\lambda$), la diffraction n'est observée que dans les directions de \mathbf{s} où le vecteur de diffraction $\mathbf{S} = \mathbf{s} - \mathbf{s}_0$ vérifie les équations de Laue :

$$\mathbf{S} \cdot \mathbf{a} = h \qquad \mathbf{S} \cdot \mathbf{b} = k \qquad \mathbf{S} \cdot \mathbf{c} = l$$

où h , k et l sont des entiers.

Les équations de Laue se transforment en :

$$\mathbf{S} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*$$

où \mathbf{S} est un vecteur du réseau réciproque ayant pour vecteur de base \mathbf{a}^* , \mathbf{b}^* et \mathbf{c}^* (voir Figure II.5.A). Le réseau réciproque est construit géométriquement à partir du réseau direct par les relations :

$$\mathbf{a}^* = \mathbf{b} \wedge \mathbf{c} / V \qquad \mathbf{b}^* = \mathbf{c} \wedge \mathbf{a} / V \qquad \mathbf{c}^* = \mathbf{a} \wedge \mathbf{b} / V$$

où $V = \mathbf{a} \cdot \mathbf{b} \wedge \mathbf{c}$ est le volume de la maille du réseau direct. Chaque réflexion correspond à un nœud du réseau réciproque et peut ainsi être indexée par un jeu de trois indices h , k , et l appelés indices de Miller.

La **loi de Bragg** formule de manière différente ces conditions, en introduisant le concept de plan réticulaire. La diffraction des rayons X est alors assimilée à une réflexion du faisceau incident sur une famille de plans réticulaires parallèles (perpendiculaires à \mathbf{S}) d'indice h, k, l et de distance interréticulaire d_{hkl} , se produisant aux angles 2θ (angle entre les faisceaux incidents et diffractés, donc entre \mathbf{s}_0 et \mathbf{s} voir Figure II.5.B) vérifiant l'équation :

$$|\mathbf{S}| = 2|\mathbf{s}_0| \sin \theta = \frac{2 \sin \theta}{\lambda} = \frac{1}{d_{hkl}}$$

A une réflexion d'indice h, k, l correspond une résolution dans l'espace direct (aussi appelée espacement de Bragg), mesurée par d_{hkl} , la distance entre plans réticulaires diffractants correspondantes. Dans le cas de l'utilisation d'une source de rayons X monochromatique, à chaque faisceau diffracté correspond une seule réflexion d'indice h, k, l .

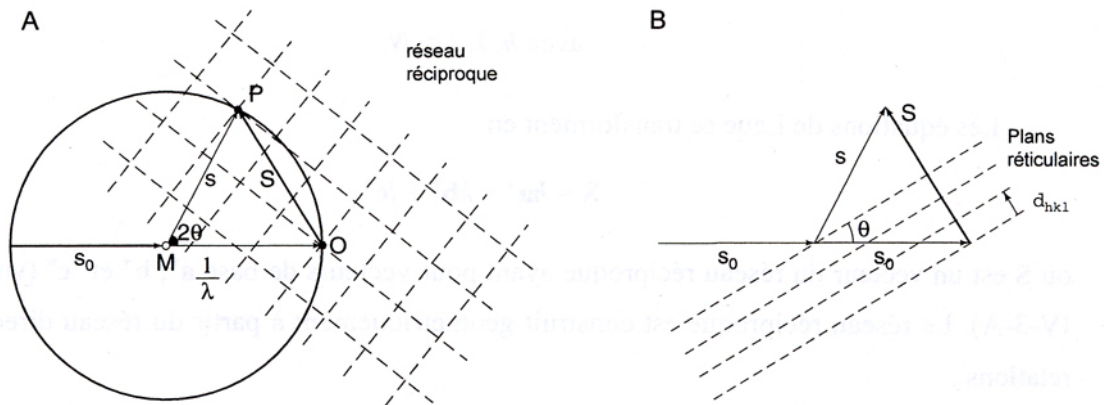


Figure II.5 : Schéma décrivant le phénomène de la diffraction selon les lois de Laue (A) et de Bragg (B).

2.1.2 Les facteurs de structure

L'onde diffusée par un objet est la somme de la diffusion de tous les atomes de l'objet et peut donc être représentée mathématiquement comme la transformée de Fourier de cet objet. Un cristal peut-être représenté comme le produit de convolution d'une fonction motif (les molécules présentes dans la maille du cristal) par une fonction réseau qui répète ce motif par translation dans le cristal (défini par les paramètres de maille). Or, la transformée de Fourier échange produit ordinaire et produit de convolution. En effet la transformée de Fourier d'un produit de convolution est égale au produit des transformées de Fourier. Les ondes diffractées correspondent donc à la transformée de Fourier de la densité électronique du motif, échantillonnée aux nœuds du réseau réciproque. Ces réflexions peuvent être représentées par un nombre complexe \mathbf{F}_{hkl} appelé facteur de structure qui peut donc être calculé de la manière suivante :

$$\mathbf{F}_{hkl} = V \int_{x=0}^1 \int_{y=0}^1 \int_{z=0}^1 \rho(x,y,z) e^{2\pi i(hx+ky+lz)} dx dy dz \quad (1)$$

ρ est la densité électronique et x,y,z sont les coordonnées fractionnaires de l'espace direct. En appliquant la transformée de Fourier inverse, qui se réduit dans ce cas à une série de Fourier (\mathbf{F}_{hkl} n'existant que pour les valeurs h, k et l entières), nous pouvons retrouver l'image de la densité électronique du motif diffractant :

$$\rho(x,y,z) = \frac{1}{V} \sum_h \sum_k \sum_l \mathbf{F}_{hkl} e^{-2\pi i(hx+ky+lz)} \quad (2)$$

Chaque facteur de structure \mathbf{F}_{hkl} peut-être défini par une amplitude F_{hkl} et une phase Φ_{hkl} . Ces deux composantes doivent être connues simultanément pour reconstruire la densité. Un modèle atomique (décrivant la nature, la position, l'agitation et l'occupation de site de N atomes) permet de calculer les facteurs de structure par la relation suivante :

$$\mathbf{F}_{hkl} = F_{hkl} e^{i\Phi_{hkl}} = \sum_{j=1}^N q_j f_j e^{-B_j \left(\frac{\sin \theta}{\lambda} \right)^2} e^{2\pi i(hx_j+ky_j+lz_j)} \quad (3)$$

où f_j , B_j , q_j , x_j , y_j et z_j sont respectivement le facteur de diffusion atomique, le facteur d'agitation thermique, l'occupation et les trois coordonnées de l'atome j . C'est l'application de l'équation (1) à un modèle atomique discontinu.

Il n'y a pas de moyen direct pour mesurer la phase Φ_{hkl} des facteurs de structures \mathbf{F}_{hkl} (voir § II.2.5). En effet, les détecteurs ne sont sensibles qu'à l'intensité I_{hkl} des ondes diffractées. Or l'intensité, grandeur réelle, ne dépend que de l'amplitude du facteur de structure :

$$I_{hkl} = \mathbf{F}_{hkl} \overline{\mathbf{F}_{hkl}} = |\mathbf{F}_{hkl}|^2 \quad (4)$$

Il n'est donc pas possible d'en déduire la phase ni de reconstruire directement la densité électronique par la seule transformée de Fourier inverse.

2.2 Enregistrement des données

La méthode classique qui est utilisée en cristallographie des macromolécules consiste à enregistrer les intensités diffractées sur des détecteurs bidimensionnels, au cours d'un ensemble de petites rotations du cristal, par pas de 0,5 ou 1° par exemple. Cette méthode permet de rapidement mettre en condition de diffraction une grande partie des réflexions de l'espace réciproque. Le temps d'exposition dépend de l'intensité du faisceau incident et du pouvoir de diffraction du cristal. Dans la grande majorité des cas un faisceau incident monochromatique est utilisé.

Les rayons X proviennent soit d'un générateur à anode tournante, soit d'un rayonnement synchrotron. Dans le premier cas, la longueur d'onde est généralement celle de la raie d'émission $K\alpha$ du cuivre, $\lambda=1,54 \text{ \AA}$, alors que dans le second, une gamme continue de longueurs d'ondes est accessible, typiquement de 0,5 à 2 \AA .

Les différents jeux de données des cristaux de la protéine IRP1 ont été collectés sur les lignes BM30a et ID14 de l'*European Synchrotron Radiation Facility* (ESRF) de Grenoble.

2.2.1 La ligne de lumière FIP – BM30a

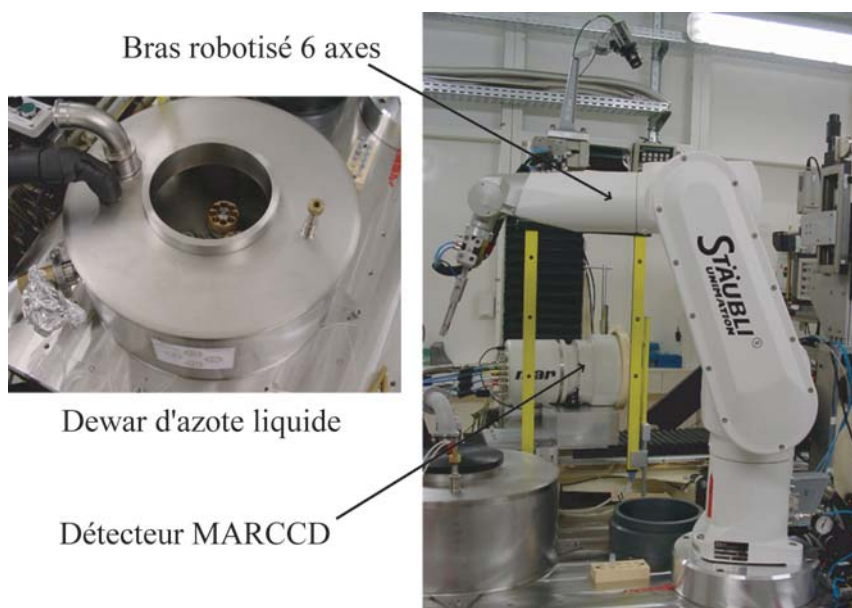


Figure II.6 : Système CATS (Cryogenic Automated Transfert System). A gauche, le Dewar rempli d'azote contenant 4 carrousels capables d'accueillir 10 échantillons chacun. A droite, le bras robotisé utilisé pour monter et récupérer les échantillons à analyser.

La ligne FIP-BM30a est une ligne permettant d'atteindre des longueurs d'onde comprises entre 0,7 et 1,8 \AA avec un faisceau de 300 μm par 300 μm . Les échantillons sont placés sur une tête goniométrique à 4 cercles et les données sont enregistrées par un détecteur MARCCD (Charged Couple Device). Pendant toute la collecte, les cristaux sont maintenus à une température de 100 K, grâce au système cryogénique « Oxford cryostream ». La particularité de la ligne FIP réside dans le fait que la majorité des différentes étapes pour mener à bien une expérience cristallographique sont automatisées (Roth, 2002). De l'optimisation de l'intensité du faisceau en passant par son centrage ainsi que le choix de la longueur d'onde utilisée (indispensable pour le phasage par les méthodes SAD ou MAD) toutes ces étapes sont complètement automatiques. De même le centrage du puit ou, en anglais, « *beam-stop* », outil qui empêche le faisceau principal de rayons X d'atteindre directement le détecteur risquant de l'endommager, est informatisé. Un système robotisé

nommé « CATS » (Cryogenic Automated Transfer System) permet le montage ainsi que la récupération des échantillons de façon automatique (Figure II.6).

L'avantage majeur de ce système est qu'il permet d'améliorer la stratégie de collecte et d'optimiser le temps d'occupation de la ligne de lumière. D'une part il permet la récupération aisée et à moindre risque d'un cristal, facilitant le choix du meilleur cristal, et d'autre part il limite les accès sécurisés à la cabine expérimentale. Le centrage du cristal est géré par un système semi-automatique. Enfin, dans les cas les plus favorables un logiciel nommé « ADP » (Automated Data Processing) permet d'analyser des données et de résoudre la structure cristallographique de la protéine d'intérêt.

2.2.2 Les lignes de lumière ID14

Les lignes ID14-1 et 2 sont pourvues d'un faisceau monochromatique dont la longueur d'onde est fixée à environ 0.93 Å. Le diamètre du faisceau est compris entre 50 et 200 µm. Ces lignes, pourvues d'un goniomètre à 1 cercle, sont équipées d'un détecteur ADSC Q4R CCD. La température du cristal est maintenue à 100 K par un système cryogénique « Oxford cryostream ». La résolution maximale qu'il est possible d'atteindre est de 0.97 Å.

Enfin les lignes ID14-3 et 4 est pourvue d'un micro-diffractomètre avec centrage du cristal semi-automatique. Le détecteur est un appareil MARCCD3x3. La longueur d'onde est variable de 0,62 à 2,07 Å.

2.3 Le traitement des données

2.3.1 Les principes

Le protocole de traitement des images enregistrées lors de la rotation du cristal comporte trois étapes essentielles :

- **l'indexation** : cette étape consiste, à l'aide de la position des tâches de diffraction couplée à l'angle d'oscillation, à déterminer l'orientation du cristal et ses paramètres de maille. Les programmes d'indexation cherchent trois vecteurs de base permettant de relier toutes les réflexions, puis la matrice d'orientation ainsi que la maille qui expliquent toutes les réflexions observées. Il est alors possible de prédire la position de l'ensemble des tâches de diffraction sur les images et d'en mesurer l'intensité ;

- **l'intégration des intensités** : la dispersion spatiale des raies de diffraction, due essentiellement à la mosaïcité du cristal, à la largeur spectrale et à la divergence du faisceau incident, a pour conséquence d'étaler une tâche sur un diamètre de quelques millimètres. Les programmes de traitements permettent d'établir les profils moyens de ces tâches et de calculer

l'intensité de chaque réflexion par intégration numérique suivant ces profils calculés. Il est aussi nécessaire de corriger les effets d'absorption et de polarisation du faisceau incident ;

- **la mise à l'échelle** : du fait de la dégradation des cristaux dans le flux de rayons X, de la décroissance de l'intensité du faisceau synchrotron ou de la forme du cristal, l'intensité des réflexions normalement identiques par symétrie et mesurée à des instants différents et pour des orientations différentes ne sera pas égale. Pour tenir compte de tous ces paramètres, il faut donc mettre au même niveau d'intensité ces différentes réflexions. Les programmes de mise à l'échelle vont déterminer deux facteurs, un linéaire, k , et l'autre en fonction de la résolution, B , pour chaque image à partir des réflexions supposées équivalentes par les symétries du cristal. Par la suite, les intensités de ces réflexions équivalentes sont moyennées et le module des facteurs de structure est calculé : c'est la réduction des données qui permet d'obtenir l'ensemble des réflexions uniques nécessaires pour calculer l'image de la densité électronique par transformée de Fourier.

2.3.2 Les programmes utilisés

Il existe toute une série de programmes dédiés au traitement des images de diffraction qui permettent d'extraire les intensités de chaque réflexion, de mettre à l'échelle et de calculer les facteurs de structure. Dans cette étude, j'ai utilisé différents programmes pour traiter les jeux de données qui ont été enregistrés : XDS (Kabsch, 1993) pour l'indexation et l'intégration des tâches de diffraction et la suite CCP4 (CCP4, 1994) pour la réduction des données.

- **le programme XDS** : ce programme, hautement automatisé, développé par W. Kabsch se décompose en sept sous-programmes qui permettent un traitement rapide des données en découplant les différentes opérations. Il a aussi l'avantage de déterminer des profils tridimensionnels indispensables dans le cas où les cristaux sont mosaïques et pour lesquels l'intensité d'une tâche peut être répartie sur plusieurs images. Par contre, ce programme reste plus difficile d'approche par son manque d'interface graphique et le peu de paramètres manipulables pour l'intégration.

- **la suite CCP4** : cette suite de programmes dispose d'un ensemble de scripts pour la manipulation des fichiers informatiques contenant les réflexions mesurées (COMBAT, SORTMTZ, CAD...) mais aussi les programmes SCALA et TRUNCATE qui ont servi à la mise à l'échelle et au calcul des modules des facteurs de structure.

2.3.3 La validation des données

La qualité d'un jeu de données est déterminée par différents critères statistiques calculés au cours du traitement. Les paramètres les plus significatifs et utilisés sont :

- **la résolution** : qui déterminera la finesse des détails dans l'image de la densité électronique. La limite haute de la résolution des données mesurées peut être imposée par la configuration de l'expérience (taille du détecteur, distance cristal-détecteur, longueur d'onde...) ou par la qualité de diffraction du cristal. Dans ce dernier cas, un rapport signal sur bruit moyen de l'ordre de 2 est communément admis comme seuil de coupure ;

- **le signal sur bruit (I/σ) moyen** : plus il est grand et meilleure est l'estimation des facteurs de structure ;

- **la complétude** : qui mesure le rapport entre le nombre de réflexions uniques mesurées et le nombre théorique de ces réflexions pour la résolution des données. Elle doit être la plus grande possible et supérieure à 90% ;

- **la redondance** : qui correspond au nombre de fois où l'intensité d'une réflexion unique est mesurée. Ce qui est équivalent au nombre moyen de réflexions symétriques mesurées. Plus la valeur numérique de ce critère est grande, meilleure sera l'estimation de l'intensité moyenne d'une réflexion unique en diminuant l'erreur commise à condition que le cristal ne se dégrade pas pendant la collecte ;

- **le facteur R_{sym}** : (ou R_{merge} si plusieurs cristaux sont utilisés) qui représente le désaccord entre les intensités des réflexions équivalentes par symétrie. Plus ce facteur est bas, meilleure est la cohérence du jeu de données. Il s'exprime par :

$$R_{\text{sym}} = \frac{\sum_{h'k'l'} \sum_{hkl} |I_{hkl} - \langle I_{h'k'l'} \rangle|}{\sum_{hkl} I_{hkl}} \quad (5)$$

où $\langle I_{h'k'l'} \rangle$ est l'intensité moyenne de la réflexion unique $h'k'l'$ et I_{hkl} est l'intensité mesurée de la réflexion hkl .

Toutefois, il a pu être montré que l'augmentation de ce facteur statistique en fonction du nombre de réflexions n'était pas seulement due à l'accumulation des erreurs systématiques lors de la mesure mais aussi à la redondance du jeu de données. (Diederichs and Karplus, 1997) ont alors proposé l'utilisation d'un terme statistique qui reflète plus justement la fiabilité de chaque réflexion indépendamment de la redondance. Ce facteur nommé R_{meas} est obtenu en pondérant chaque réflexion par un terme dérivé de la redondance :

$$R_{\text{meas}} = \frac{\sum_{h'k'l'} \sqrt{\frac{n_{h'k'l'}}{n_{h'k'l'} - 1}} \sum_{hkl} |I_{hkl} - \langle I_{h'k'l'} \rangle|}{\sum_{hkl} I_{hkl}} \quad (6)$$

où $n_{h'k'l'}$ est la redondance de la réflexion unique $h'k'l'$.

• **la distribution des intensités en fonction de la résolution** : par son analyse statistique, il est parfois possible de déterminer si le cristal est maclé et de calculer la proportion de chaque partie de la macle. Les relations entre les réflexions, qu'imposent les lois de macle, et la proportion de chaque partie diffractante donne lieu à des distributions des intensités particulières qui sont caractérisées par leurs moments dont les valeurs sont tabulées (Yeates, 1997).

2.4 Caractéristiques des cristaux de protéines

2.4.1 Régularité du réseau cristallin ou mosaïcité

Les cristaux de protéines présentent souvent des contacts intermoléculaires assez réduits. Aussi, des perturbations mécaniques ou thermiques intervenant lors de leur croissance ou de leur manipulation peuvent engendrer des défauts dans leur réseau cristallin. La qualité de ce réseau peut être caractérisée par un facteur appelé « mosaïcité » et exprimé en degré. Typiquement, il varie de 0,2 à 1,5° et exprime l'imperfection des cristaux qui sont souvent composés d'une mosaïque de pavés cristallins plus ou moins légèrement désorientés. C'est donc la différence d'orientation moyenne entre chaque bloc qui est mesurée. Expérimentalement, plus la mosaïcité est grande et plus la divergence des faisceaux diffractés est importante. Une forte mosaïcité risque d'entraîner des problèmes de superposition de réflexions, ou de causer des difficultés pour mesurer les réflexions de faibles intensités.

2.4.2 Désordre des atomes du cristal

Les macromolécules biologiques ne sont pas des structures figées. Certaines régions de ces molécules, en particulier celles qui sont exposées au solvant peuvent présenter une certaine variabilité ou agitation conformationnelle. Dans un cristal, ce désordre est d'autant plus marqué que la température est élevée et le cristal faiblement compact. Il se traduit par un affaiblissement global des intensités diffractées, qui peut être pris en compte par un facteur d'agitation thermique global (voir équation 3) exprimé sous la forme $\exp[-B(\sin\theta/\lambda)^2]$. De plus, certaines régions de ces molécules et en particulier celles qui sont exposées au solvant,

peuvent présenter une plus grande variabilité dans leur position et ainsi participer d'une moindre façon à la diffraction des rayons X. Il est alors difficile d'observer la densité électronique de ces régions désordonnées. Le facteur B moyen est exprimé en \AA^2 , et reflète aussi bien le désordre statique que le dynamique. Généralement, les cristaux de protéines possèdent des facteurs B moyens compris entre 10 et 70 \AA^2 .

2.5 Analyse du contenu de l'unité asymétrique

Il est possible de décrire la symétrie des cristaux de macromolécules biologiques par 64 types de symétries distinctes appelées groupes d'espaces (Tableau II.1).

Triclinic None	$P1$
Monoclinic Unique axis b : $\alpha = \gamma = 90^\circ$	$P2, P2_1, C2$
Orthorhombic $\alpha = \beta = \gamma = 90^\circ$	$P222, P2_12_12_1, P2_12_12$ $P222_1, C222, C222_1, F222$ $I222, I2_12_12_1$
Tetragonal $a = b; \alpha = \beta = \gamma = 90^\circ$	$P4, P4_1, P4_2, P4_3, I4, I4_1$ $P422, P4_22, P4_122$ $P4_12_12, P4_222, P4_22_12$ $P4_32_12, P4_322, I4_122$
Trigonal and rhombohedral Hexagonal axes: $a = b; \alpha = \beta = 90^\circ; \gamma = 120^\circ$ Rhombohedral axes: $a = b = c; \alpha = \beta = \gamma \neq 90^\circ$	$P3, P3_1, P3_2, R3$ $P312, P321, P3_121, P3_112$ $P3_221, P3_212, R32$
Hexagonal $a = b; \alpha = \beta = 90^\circ; \gamma = 120^\circ$	$P6, P6_5, P6_1, P6_4, P6_3$ $P6_2, P622, P6_122, P6_222$ $P6_322, P6_422, P6_522$
Cubic $a = b = c; \alpha = \beta = \gamma = 90^\circ$	$P23, F23, I23, P2_13, I2_13$ $P432, P4_132, P4_232, P4_332$ $P432, F4_132, I432, I4_132$

Tableau II.1 : Systèmes cristallins et groupes d'espace.

L'unité asymétrique de la maille est le plus petit volume de la maille contenant des points non reliés entre eux par la symétrie cristallographique. Autrement dit, l'ensemble de la maille est engendré par l'unité asymétrique à laquelle il est appliqué les opérateurs de symétrie du groupe d'espace (Figure II.7). Elle peut contenir plusieurs molécules qui sont alors reliées entre elles par des opérateurs de rotation - translation définissant une symétrie non cristallographique (SNC). Nous exposerons dans ce paragraphe les méthodes qui permettent d'estimer ou de mesurer le nombre de molécules par unité asymétrique et d'identifier dans les cas simples la SNC qui les relie.

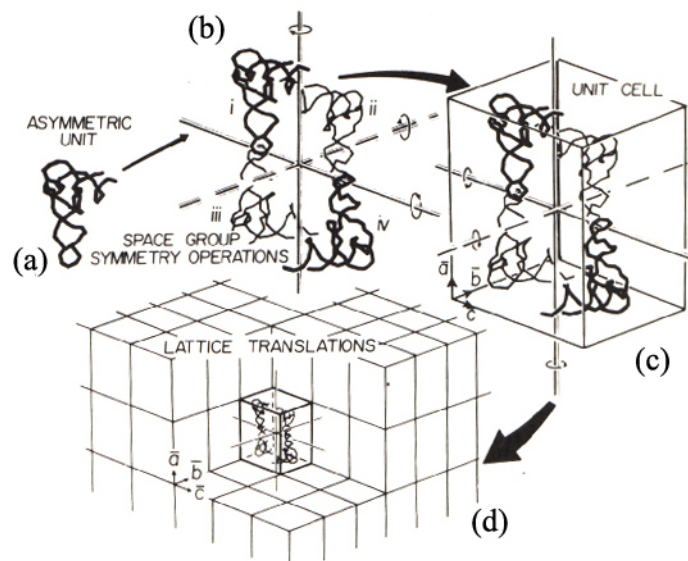


Figure II.7 : Cristal généré à partir de l'unité asymétrique. Cet exemple représente le squelette d'un ARN de transfert de levure (a), en lui appliquant les opérateurs de symétrie de son groupe d'espace (ici trois axes perpendiculaires d'ordre 2) (b) tout en choisissant le système cristallin qui contient l'ensemble de l'unité asymétrique (ici un système orthorhombique)(c) et en répétant cette maille et son contenu de manière périodique le long des trois directions définies par les axes de l'unité cellulaire nous aboutissons à la reconstitution du cristal (d). D'après McPherson (1989).

2.5.1 Estimation du nombre de molécules de l'unité asymétrique (Z_a)

Les cristaux de protéine contiennent une forte proportion de solvant qui peut varier de 30 à 80% du contenu de la maille.

A partir du poids moléculaire de la protéine cristallisée et des paramètres de l'arrangement cristallin (maille et groupe d'espace), il est possible d'estimer la quantité de solvant dans le cristal ($S_{\%}$) ainsi que le nombre de molécules de protéine contenues dans l'unité asymétrique (Z_a). Cette analyse repose sur une méthode proposée par Matthews (1968) : il a montré que, pour la plupart des cristaux protéiques, le rapport (V_m) entre le volume de la maille (V) et le poids moléculaire contenu dans celle-ci était compris entre 2 et 4 $\text{\AA}^3/\text{Da}$. Le coefficient V_m est déterminé par :

$$V_m = \frac{V}{M \cdot n \cdot Z_a} \quad (7)$$

où V est le volume de la maille, n le nombre de positions équivalentes, Z_a le nombre de molécules dans l'unité asymétrique et M la masse moléculaire de la protéine. La quantité de solvant peut quant à elle être estimée comme :

$$S_{\%} = 1 - \frac{1,23}{V_m} \quad (8)$$

où 1,23 g/cm^3 est la valeur de l'inverse de la masse spécifique moyenne des protéines (0,807 Da/\AA^3).

2.5.2 Identification des symétries non cristallographiques

Lorsque plusieurs molécules sont présentes à l'intérieur de l'unité asymétrique ($Z_a > 1$), leurs positions respectives sont décrites par des relations de symétrie non cristallographique (SNC). La SNC est qualifiée de « symétrie propre » quand elle peut être réduite à une symétrie simple (un axe de rotation) ou de « symétrie impropre » dans le cas d'une symétrie plus complexe (la combinaison d'un axe de rotation et d'une translation). Parfois il est possible d'identifier la SNC avant d'avoir résolu le problème du « phasage », grâce à la fonction de Patterson : $P(u,v,w)$. Cette fonction est le produit de convolution de la densité électronique $\rho(x,y,z)$ contenue dans la maille avec son inverse par rapport à l'origine, mais c'est aussi une fonction d'auto-corrélation de la densité :

$$P(u,v,w) = \frac{1}{V} \int \int \int \rho(x,y,z) \rho(u+x,v+y,w+z) dx dy dz = \frac{1}{V} \rho(x,y,z) \times \rho(-x,-y,-z) \quad (9)$$

Par application de l'équation (2), cette fonction peut alors se calculer, par synthèse de Fourier, à partir des intensités diffractées grâce à la relation :

$$P(u,v,w) = TF(F_{hkl}) \times TF(\overline{F}_{hkl}) = TF(F_{hkl} \overline{F}_{hkl}) = TF(I_{hkl}) \quad (10)$$

ce qui peut se ramener à :

$$P(u,v,w) = \frac{1}{V} \sum_{h,k,l} I_{hkl} \cos(2\pi(hu + kv + lw)) \quad (11)$$

Cette fonction possède des valeurs élevées pour toute distance de corrélation de la densité $t(u,v,w)$ qui fait se correspondre des régions de forte densité. Ainsi, l'existence d'une SNC de translation simple sera révélée dans cette fonction par un pic très intense ayant les coordonnées du vecteur de translation. Pour une SNC de rotation, la direction et l'ordre de l'axe de rotation (mais pas sa position) peuvent être détectés par la fonction de « *self-rotation* » (Rossmann et Blow, 1962) qui est la fonction d'auto-corrélation en rotation de la fonction de Patterson. Si, pour une rotation, ne correspondant pas à une rotation cristallographique, la corrélation est forte, cela met en évidence une relation de SNC de rotation entre des molécules de l'unité asymétrique. La fonction de self-rotation révèle les éléments de SNC de rotation propres et impropres (elle est insensible aux translations). La fonction de Patterson fait apparaître les vecteurs de translation des SNC propres (elle est sensible aux rotations). Remarquons cependant que ces deux méthodes peuvent aussi mettre en évidence des pseudo-symétries internes aux molécules, qui peuvent être difficiles à distinguer de la SNC à *a priori*.

2.6 Méthodes de phasage

2.6.1 Le problème des phases

Le deuxième obstacle de la cristallographie, après celui de l'obtention de cristaux de bonne qualité, vient de l'incapacité à mesurer directement l'information de déphasage des ondes diffractées, les unes par rapport aux autres. En effet, il n'existe pas actuellement de méthode expérimentale permettant d'enregistrer les phases ou de systèmes optiques permettant de former une image à partir des faisceaux de rayons X durs diffractés (comme cela peut être réalisée en microscopie électronique). Les détecteurs bidimensionnels ne sont sensibles qu'à l'intensité des faisceaux diffractés. Or, pour construire l'image de la densité diffractante par transformée de Fourier, il est indispensable de posséder ces deux informations à la fois. Des essais numériques simples montrent que, lors de la reconstruction, la contribution des phases est beaucoup plus importante que celle des intensités.

Alors que dans le cas de la cristallographie des petites molécules ce problème est résolu de manière automatique (grâce aux « méthodes directes »), en ce qui concerne les macromolécules, en revanche, ce n'est pas aussi facile de surmonter cet obstacle. Néanmoins, ces dernières années les méthodes permettant de résoudre ce problème ont réalisé de nombreux progrès tant sur le plan de l'approche expérimentale que sur celui du traitement numérique des données (Carter et Sweet, 1997). Elles peuvent être classées en deux catégories, selon qu'elles utilisent comme source de phase un modèle atomique homologue (méthode du remplacement moléculaire) ou le signal apporté par des atomes « lourds ».

Une fois un premier jeu de phases déterminé, des méthodes de modification de la densité peuvent être utilisées pour en améliorer la qualité. L'amélioration des phases se poursuit ensuite en même temps que la construction puis l'amélioration du modèle. Dans tous les cas, le succès de toutes ces méthodes est conditionné par la qualité de diffraction des cristaux.

2.6.2 Remplacement moléculaire

L'utilisation de la méthode du remplacement moléculaire (Rossmann, 1972) s'est significativement développée ces dernières années. Son intérêt s'accroît, dans la mesure où le nombre de structures pouvant être utilisées comme modèles augmente rapidement. Le but est de positionner une ou plusieurs molécules modèles dans la maille du cristal étudié afin de pouvoir calculer une première estimation des phases pour les facteurs de structure. Le

problème consiste dans ce cas à trouver la matrice de rotation et le vecteur de translation permettant de simuler le plus précisément possible le contenu de la maille étudiée. La recherche des fonctions de rotation (RF) et de translation (TF) se fait grâce à des fonctions de corrélation permettant de comparer les amplitudes mesurées (F_o) et les amplitudes calculées à partir du modèle (F_c).

2.6.2.1 Fonction de rotation

La première étape de la méthode du remplacement moléculaire est la recherche de l'orientation du modèle initial. La méthode de M. Rossmann et D. Blow (Rossmann and Blow, 1962) consiste à déterminer les trois angles d'Euler ($\theta_1, \theta_2, \theta_3$) pour lesquels la fonction de corrélation entre la fonction de Patterson du modèle orienté et celle issue du jeu de vecteurs observés est maximale.

$$RF(\theta_1, \theta_2, \theta_3) = \int_V P_{\text{obs}}(\mathbf{u}) \cdot U(\mathbf{u}) \cdot P_{\text{calc}}(\mathbf{R} \cdot \mathbf{u}) d\mathbf{u} \quad (12)$$

où $P_{\text{obs}}(\mathbf{u})$ est la fonction de Patterson observée en position \mathbf{u} , \mathbf{R} est la matrice de rotation correspondant aux angles $\theta_1, \theta_2, \theta_3$ et $P_{\text{calc}}(\mathbf{R} \cdot \mathbf{u})$ est la fonction de Patterson calculée en position $\mathbf{R} \cdot \mathbf{u}$. Elle est équivalente à la fonction de Patterson calculée en \mathbf{u} ayant subi la rotation \mathbf{R} , $U(\mathbf{u})$ est une fonction masque limitant le volume d'intégration.

Le calcul de cette fonction de corrélation doit être limité à la distribution des vecteurs entre atomes d'une seule molécule. Le volume d'intégration est choisi de telle sorte que le maximum de vecteurs intra-moléculaires soit inclus tout en excluant un maximum de vecteurs inter-moléculaires. Il est typiquement de l'ordre du rayon du modèle utilisé pour la recherche en rotation.

2.6.2.2 Fonction de translation

Durant la seconde étape du remplacement moléculaire, le modèle correctement orienté doit être positionné en translation dans la nouvelle maille. La méthode la plus rapide numériquement est celle développée par R. Crowther et D. Blow (Crowther and Blow, 1967) qui utilise les vecteurs entre molécules symétriques de la fonction de Patterson. Si le modèle est placé de façon arbitraire dans la maille, la position des symétriques est tout aussi arbitraire et détermine une distribution singulière des vecteurs inter-moléculaires. Ainsi, la fonction de translation définie par Crowther et Blow, comme la corrélation entre les distributions des vecteurs et ceux issus des facteurs de structure observés sera maximale pour un vecteur \mathbf{t} reliant deux positions symétriques réelles.

$$TF(\mathbf{t}) = \int_V P_{\text{obs}}(\mathbf{u}) \cdot P_{\text{calc}}(\mathbf{u}, \mathbf{t}) d\mathbf{u} \quad (13)$$

où $P_{\text{obs}}(\mathbf{u})$ est la fonction de Patterson observée en position \mathbf{u} et $P_{\text{calc}}(\mathbf{u}, \mathbf{t})$ est la fonction de Patterson calculée en position \mathbf{u} et paramétrée en fonction de \mathbf{t} .

2.6.2.3 Les conditions

Cette méthode de phasage requiert certaines conditions qui doivent être respectées pour garantir une solution de remplacement moléculaire correcte :

- **un jeu de données le plus complet possible** et de bonne qualité dans toute la gamme de résolution utilisée ;
- **un modèle structural le plus proche possible**. La comparaison des séquences en acides aminés est une bonne indication de l'homologie structurale. Deux molécules ayant plus de 30 % d'identité séquentielle ont toutes les chances d'avoir un repliement similaire. De plus, il peut aussi être préférable de modifier le modèle en supprimant les chaînes latérales ou certaines parties (boucles, domaines) lorsqu'elles sont supposées trop éloignées de la structure de la protéine inconnue ;
- **une gamme de résolution appropriée**. Il est préférable d'exclure les données à haute résolution lorsque les structures sont supposées sensiblement éloignées. En général, une gamme de résolution de 15 à 3.5 Å est utilisée ;
- **un rayon d'intégration** séparant au mieux les vecteurs intra- et inter-moléculaires.

2.6.3 Utilisation d'atomes lourds

Les techniques les plus anciennement employées utilisent les propriétés particulières de diffraction associées à la présence d'atomes dits lourds (à numéro atomique élevé par rapport à ceux des atomes de la protéine), soit naturellement présents dans la structure, soit introduits dans le cristal à cet effet.

Les premières méthodes ayant été développées dans ce sens sont les méthodes de remplacement isomorphe (MIR et SIR) qui exploitent la contribution de ces atomes lourds à la diffraction. Il est aussi possible de tirer parti d'un phénomène physique résultant de la diffusion d'un rayonnement X par des atomes : la diffusion anormale qui est maximale à leurs seuils d'absorption.

2.6.3.1 Méthode MIR

La méthode MIR a été introduite par l'équipe de M. Perutz, pour résoudre la structure de l'hémoglobine (Green et al., 1954).

Le principe de la méthode consiste à mesurer les spectres de diffraction de cristaux de la protéine marquée aux atomes lourds. Les atomes en question doivent posséder un nombre d'électrons notablement plus grand que celui des atomes de la biomolécule (hydrogène, carbone, azote, oxygène et soufre), afin que l'intensité des taches de diffraction des cristaux contenant des atomes lourds fixés à la protéine (cristaux dérivés) diffère notablement de celle des taches de diffraction correspondant aux cristaux de la protéine seule (cristaux natifs). Il est ainsi possible d'estimer la phase de chaque réflexion à partir des différentes intensités observées.

Comme nous le verrons ultérieurement, plusieurs dérivés d'atomes lourds sont nécessaires pour déterminer la phase. L'une des conditions essentielles de la réussite de la méthode est que la structure des cristaux dérivés soit très similaire à celle du cristal natif : les paramètres de maille et l'orientation des molécules dans la maille doivent être identiques ou presque pour les deux cristaux.

Cet isomorphisme permet de traduire la modification due à la présence de l'atome lourd par égalité vectorielle :

$$\mathbf{F}_{PH} = \mathbf{F}_P + \mathbf{F}_H \quad (14)$$

où \mathbf{F}_{PH} , \mathbf{F}_P et \mathbf{F}_H représentent, respectivement, le facteur de structure du cristal dérivé, celui de la protéine seule et celui de l'atome lourd, pour une même réflexion (Figure II.8).

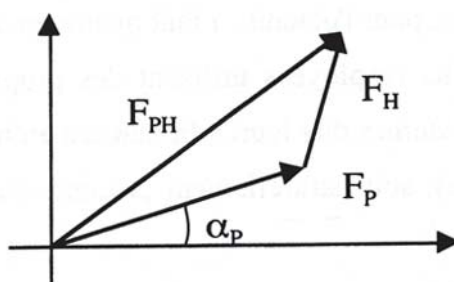


Figure II.8 : Représentation du facteur de structure de l'ensemble protéine-atome lourd. Il s'agit du résultat de la somme vectorielle du facteur de structure de la protéine et du facteur de structure de la sous-structure des atomes lourds.

Si la condition d'isomorphisme est respectée, il est possible de déterminer, par une méthode vectorielle, la phase inconnue du vecteur \mathbf{F}_P (Figure II.9).

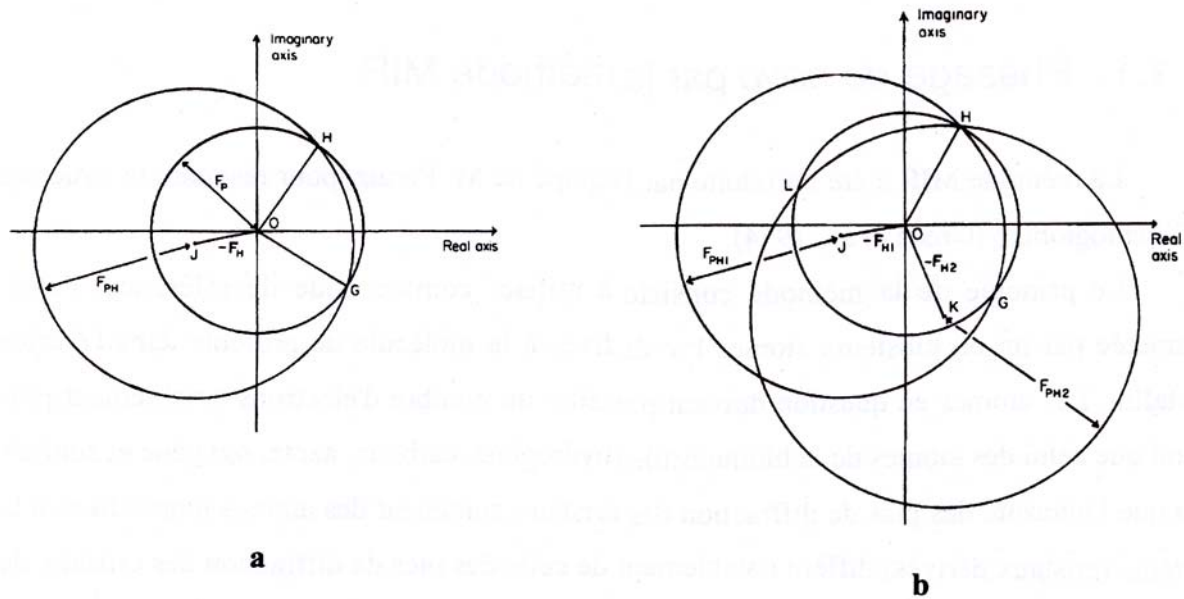


Figure II.9 : (a) Construction de Harker pour déterminer la phase par la méthode SIR (un seul dérivé lourd). Il y a deux solutions possible pour \mathbf{F}_P (les vecteurs OG et OH). (b) La méthode MIR utilise au moins deux dérivés lourds pour lever l'ambiguïté sur la phase de \mathbf{F}_P (vecteur OH solution unique). D'après Blundell et Johnson (1976).

Pour la méthode SIR, un seul dérivé est utilisé ainsi qu'un seul cristal natif. Deux cercles correspondent à ces jeux de données, pour une réflexion hkl donnée (Figure II.9.a) :

- l'un centré en O et de rayon égal au module de \mathbf{F}_P (obtenu à partir de la mesure des intensités de diffraction du cristal natif) ;
- l'autre centré à l'extrémité du vecteur $-\mathbf{F}_H$ (donné par la sous-structure des atomes lourds présents dans le cristal) et de rayon égal au module de \mathbf{F}_{PH} (mesure des intensités de diffraction du cristal dérivé). La sous-structure des atomes lourds est déterminée à l'aide de la carte de différence de Patterson entre les amplitudes des ondes diffractées par le cristal natif et par le cristal dérivé $((\mathbf{F}_{PH}-\mathbf{F}_P)^2 = \Delta \mathbf{F}_{iso}^2)$.

Les deux cercles se coupent en H et G, déterminant ainsi deux phases possibles pour le vecteur \mathbf{F}_P . L'indétermination est levée grâce à la mesure des intensités de diffraction d'un deuxième cristal dérivé (MIR). Sur la Figure II.9.b, le vecteur OH, qui correspond à l'ensemble des mesures, est donné par l'intersection des cercles représentant les trois séries de mesures.

Ces trois cercles présentent le plus souvent des rayons voisins et de plus les différents F sont mesurés avec une certaine incertitude, ce qui aboutit à une indétermination importante sur la phase. Dans ce cas, la qualité du phasage obtenu est liée à la précision des mesures. Dans la pratique, il est courant d'utiliser trois ou quatre dérivés pour déterminer plus précisément la phase.

2.6.3.1.1 Atomes lourds utilisés et modes de fixation

Les cristaux de protéines contenant des atomes lourds sont le plus souvent obtenus à partir de cristaux natifs trempés ou co-cristallisés dans une solution contenant l'atome lourd d'intérêt. Pour favoriser la fixation de l'atome lourd sur la protéine, ainsi que pour préserver l'ordre cristallin, des facteurs tels que le pH, la concentration de la solution en atome lourd ainsi que le temps de trempage sont déterminants.

Les composés les plus fréquemment utilisés sont les dérivés de platine, d'or, de mercure et d'uranium (Islam et al., 1998). Les trois premiers (ainsi que les composés d'argent, de palladium, d'iridium, d'osmium et de cadmium) forment en général des liaisons fortes de nature partiellement covalente avec les cystéines et les histidines à la surface de la protéine. L'uranium (ainsi que les lanthanides, le thallium et le plomb) ont plutôt tendance à former des liaisons de nature électrostatique avec les groupements carboxylates accessibles de la protéine. Le cas du xénon peut aussi être cité car il occupe les poches hydrophobes formées par le repliement de la protéine (Schiltz et al., 1997).

Les atomes lourds naturellement présents dans une macromolécule biologique peuvent parfois être substitués par des atomes plus lourds ayant une sphère de coordination et un rayon ionique similaires. Ainsi les lanthanides peuvent remplacer les ions magnésium ou calcium présents dans la structure tridimensionnelle de nombreuses protéines.

Grâce aux progrès techniques réalisés pour améliorer la qualité des données de diffraction (développement du rayonnement synchrotron, emploi de nouveaux détecteurs...), une nouvelle méthode, faisant appel au phénomène de diffusion anormale, a vu le jour dans les années 80.

2.6.3.2 Méthode MAD

D'abord méthode complémentaire des méthodes de remplacement isomorphe, la méthode MAD est peu à peu devenue une méthode de détermination de structure à part entière grâce, notamment, au développement du rayonnement synchrotron. Elle est actuellement l'une des méthodes les plus utilisées pour déterminer la phase des facteurs de structure des macromolécules biologiques. En effet, l'un des avantages de la méthode MAD est qu'il est possible d'obtenir plusieurs jeux de données, nécessaire à la résolution de la structure, à partir d'un seul cristal. Le problème de non-isomorphisme rencontré avec la méthode MIR est ainsi évité.

Cette méthode utilise la modification des intensités diffusées lorsque la longueur d'onde varie autour du seuil d'absorption des atomes lourds présents dans le cristal. Une contribution supplémentaire s'ajoute alors au facteur de diffusion « normal » (loin du seuil d'absorption), c'est le phénomène de diffusion anormale.

2.6.3.2.1 La diffusion anormale

2.6.3.2.1.1 Principe

La théorie classique de la diffusion cohérente d'une onde électromagnétique par un électron, ou diffusion de Thomson, suppose que l'électron est libre. Chacun des électrons de l'atome est alors accéléré sous l'action directe du champ incident et réémet, avec la même fréquence, en opposition de phase par rapport au rayonnement incident. Cependant, lorsque la longueur d'onde incidente est proche du seuil d'absorption de l'atome, une résonance s'établit et modifie la diffusion élastique, il s'agit du phénomène de diffusion anormale. La résonance modifie l'amplitude du facteur de diffusion et introduit un déphasage supplémentaire par rapport à l'onde incidente. Le facteur de diffusion de l'atome cesse d'être réel et dépend de la longueur d'onde λ :

$$^{\lambda}f = {}^0f + {}^{\lambda}f' + i{}^{\lambda}f'' = {}^0f + {}^0f \left(\frac{{}^{\lambda}f'}{{}^0f} + i \frac{{}^{\lambda}f''}{{}^0f} \right) \quad (15)$$

où 0f est la partie du facteur de diffusion de l'atome qui ne dépend pas de la longueur d'onde, ${}^{\lambda}f'$ et ${}^{\lambda}f''$ sont respectivement les parties réelle et imaginaire de la contribution anormale au facteur de diffusion, contribution qui varie fortement avec λ au voisinage du seuil d'absorption. La Figure II.10 montre les variations de f' et f'' en fonction de la longueur d'onde.

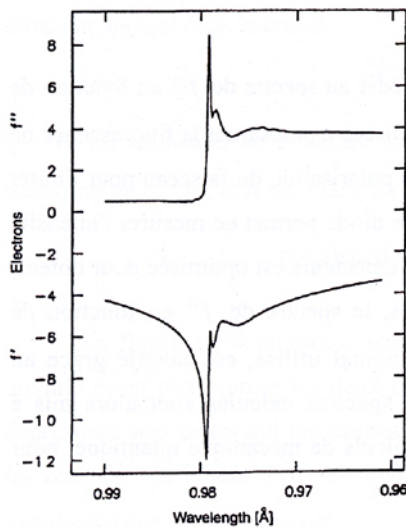


Figure II.10 : Facteurs anomaux près du seuil d'absorption K du sélénium ($\lambda=0,98$ Å) dans un cristal de thiorédoxine sélénée de *E. coli*. Il est à noter que f' redevient faible aux longueurs d'onde supérieures à celle du seuil (Hendrickson et Ogata, 1997)

Il est bon de distinguer les seuils d'absorption K, L et M qui correspondent respectivement aux transitions des orbitales 1s, 2p et 3d vers le continuum. Etant donnée la multiplicité des états d'énergie des orbitales 2p et 3d et le couplage de ces états d'énergie, il y a une succession de trois seuils L et de cinq seuils M. La contribution anormale au facteur de diffusion atomique est plus importante pour les seuils M que pour les seuils L, cette dernière étant elle-même plus importante que celle observée aux seuils K.

2.6.3.2.1.2 Obtention du facteur de diffusion anormale

La partie imaginaire du facteur de diffusion atomique f'' est liée au coefficient d'absorption μ par :

$$f''(\lambda) = K \frac{\mu(\lambda)}{\lambda} \quad (16)$$

où λ est la longueur d'onde du rayonnement utilisé et K , le coefficient de proportionnalité.

Ainsi le spectre de f'' en fonction de la longueur d'onde est le plus souvent obtenu en mesurant la fluorescence de l'échantillon aux rayons X. La fluorescence X correspond à la désexcitation radiative des électrons des couches supérieures vers les orbitales 1s, 2p et 3d, seuil K, L et M, laissées vacantes par l'absorption des photons X. Le détecteur est placé perpendiculairement à la direction de polarisation du faisceau pour limiter la diffusion élastique et donc la saturation du détecteur. Par ailleurs, une diode, ou une chambre à ionisation, permet de mesurer l'intensité du faisceau incident dans le temps. A partir de ces mesures, le spectre de f'' en fonction de l'énergie, autour du seuil d'absorption du diffuseur anormal utilisé, est calculé grâce au programme CHOOCH (Evans et Pettifer, 2001). Les spectres calculés sont alors mis à l'échelle par rapport aux valeurs obtenues à partir de calculs de mécanique quantique, pour des énergies éloignées du seuil (Cromer et Liberman, 1970).

La relation de dispersion de Kramers-Kronig permet de calculer le spectre de f' en fonction de l'énergie à partir du f'' précédemment déterminé :

$$f'(E) = \frac{2}{\pi} \int_0^{\infty} \frac{E' f''(E') dE'}{E^2 - E'^2} \quad (17)$$

Ces mesures sont nécessaires avant toute expérience MAD. En effet, cette méthode nécessite de connaître avec précision l'énergie du seuil d'absorption utilisé ainsi que l'évolution de la variation anormale dans le seuil. Cette dernière, caractéristique de la structure électronique de l'atome résonnant, dépend fortement de son environnement chimique et de son état d'oxydation (Figure II.11). Elle doit donc être mesurée sur le cristal étudié.

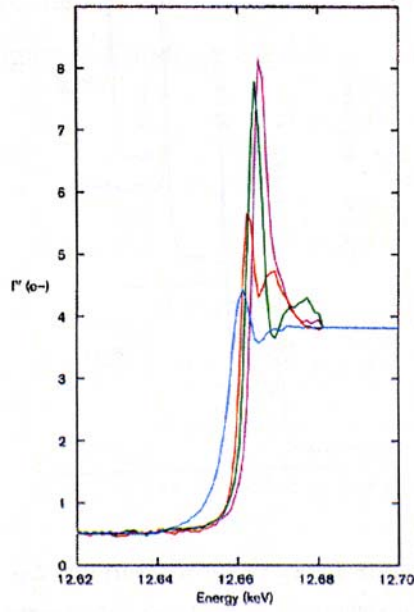


Figure II.11 : Variation de la partie imaginaire f'' du facteur de diffusion atomique pour différents complexes sélénium, au seuil d'absorption K. En bleu, la forme métallique du sélénium ; en rouge, une solution de sélénométhionine ; en vert, une solution de sélénométhionine oxydée et en rose, du sélénate (Na_2SeO_4) sous forme solide. D'après Smith et Thompson (1998).

2.6.3.2.2 Aspects théoriques de la méthode MAD

L'introduction d'un terme anomal dans le facteur de diffusion atomique (voir équation 3) induit une dépendance en fonction de la longueur d'onde pour le facteur de structure total. Le facteur de structure peut alors se décomposer en deux termes, dans le cas où il n'existe qu'un seul diffuseur anomal dans le cristal :

$${}^{\lambda}F_T(h) = {}^0F_T(h) + {}^{\lambda}F_A(h) \quad (18)$$

où 0F_T est le facteur de structure total indépendant de la longueur d'onde, loin du seuil d'absorption, et ${}^{\lambda}F_A$, le facteur de structure résonant. D'après l'équation (15), il est possible d'écrire :

$${}^{\lambda}F_T(h) = {}^0F_T(h) + \left(\frac{{}^{\lambda}f'}{{}^0f} + i \frac{{}^{\lambda}f''}}{{}^0f} \right) {}^0F_A(h) \quad (19)$$

La Figure II.12 montre comment la diffusion anormale introduit une différence en module et en phase entre les deux réflexions d'une paire de Bijvoet. $F(h)$ et $F(-h)$ sont les facteurs de structures correspondant aux deux membres d'une paire de Bijvoet.

Il suffit de multiplier ${}^{\lambda}F(\pm h)$ par son complexe conjugué pour obtenir l'expression suivante :

$$|{}^{\lambda}F(\pm h)|^2 = |{}^0F_T|^2 + a(\lambda) |{}^0F_A|^2 + b(\lambda) |{}^0F_T| |{}^0F_A| \cos({}^0\varphi_T - {}^0\varphi_A) \pm c(\lambda) |{}^0F_T| |{}^0F_A| \sin({}^0\varphi_T - {}^0\varphi_A) \quad (20)$$

où $a(\lambda)$, $b(\lambda)$ et $c(\lambda)$ sont des quantités qui dépendent des parties réelle et imaginaire du facteur de diffusion de l'atome lourd :

$$\begin{aligned} a(\lambda) &= \frac{f'^2 + f''^2}{f^2} \\ b(\lambda) &= 2 \frac{f'}{f} \\ c(\lambda) &= 2 \frac{f''}{f} \end{aligned} \quad (21)$$

Pour chaque longueur d'onde de mesure, deux équations (du type de l'équation 20) sont obtenues, correspondant aux deux réflexions d'une paire de Bijvoet. En théorie, deux longueurs d'onde (4 équations) suffisent pour résoudre ces équations (les inconnues étant les modules des facteurs de structure total $|^0F_T|$ et anomal $|^0F_A|$, ainsi que la différence des phases $\Delta\varphi = \varphi_T - \varphi_A$). En pratique, cependant, trois ou quatre jeux de données sont habituellement enregistrés à des longueurs d'onde différentes.

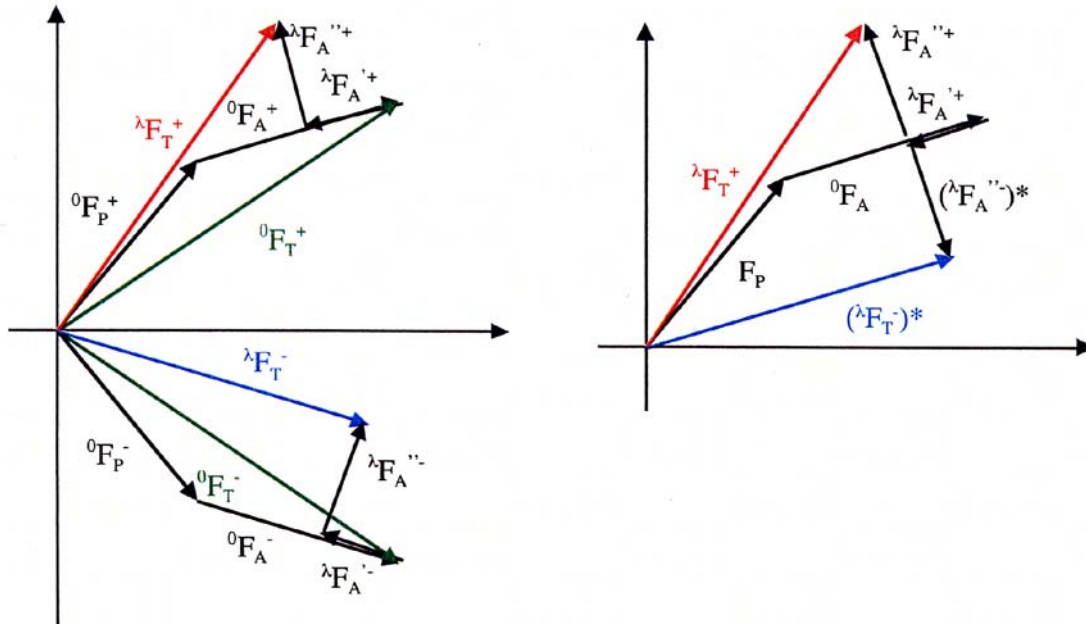


Figure II.12 : Représentation vectorielle montrant le non respect de la loi de Friedel en présence d'un diffuseur anomal. La composante imaginaire introduite par ce dernier, $^{\lambda}F_A''$, est en avance de $+\pi/2$; ce qui induit une différence d'amplitude entre $^{\lambda}F_T^+$ (en rouge) et $^{\lambda}F_T^-$ (en bleu). La différence entre les deux vecteurs $^{\lambda}F_T^+$ et $(^{\lambda}F_T^-)^*$ est équivalente à $2^{\lambda}F_A''$ (figure de droite).

La connaissance du module du facteur de structure résonnant $|^0F_A|$ donne accès, par des méthodes directes ou par le calcul de la fonction de Patterson, à la structure de référence des diffuseurs anomaux au sein du cristal (peu nombreux par rapport aux atomes de la protéine).

La phase de la structure recherchée est alors obtenue par :

$$^0\varphi_T = \Delta\varphi + ^0\varphi_A \quad (22)$$

La phase est estimée grâce à des affinements successifs des paramètres du problème : $|^0F_T|$, $|^0F_A|$, $\Delta\varphi = ^0\varphi_T - ^0\varphi_A$, ainsi que f' et f'' . Ces affinements mettent en jeu des méthodes impliquant le calcul de distributions de probabilité.

Le problème peut également être envisagé et résolu sous la forme d'une expérience MIR particulière (Ramakrishnan et Biou, 1996). Ainsi, des programmes permettant d'affiner la sous structure des atomes lourds ont pu être utilisés pour obtenir des phases à partir de jeux MAD : MLPHARE (Otwinowski, 1991) et CNS (Brünger et al., 1998) entre autres. L'utilisation du maximum de vraisemblance s'est également généralisé, notamment avec le programme SHARP (de la Fortelle et Bricogne, 1997) qui peut traiter des données MIR et MAD.

2.6.3.2.4 Protocole de mesures

2.6.3.2.4.1 Rapport anomal et « dispersif »

La diffusion anormale produit deux effets sur les intensités observées :

- D'une part, la loi de Friedel entre les deux membres d'une paire de Bijvoet n'est plus respectée. Dans le but de quantifier cet effet, le rapport de Bijvoet, ou rapport anomal, pour une longueur d'onde donnée, est défini comme :

$$\frac{\left| {}^\lambda F(h) \right| - \left| {}^\lambda F(\bar{h}) \right|}{\left\langle \left| \tilde{F} \right| \right\rangle} = \frac{\Delta F_{ano}}{F_{moyen}} \quad \text{où} \quad \left\langle \left| \tilde{F} \right| \right\rangle = \frac{\left| {}^\lambda F(h) \right| + \left| {}^\lambda F(\bar{h}) \right|}{2} \quad (23) \text{ et } (24)$$

- D'autre part, l'amplitude du facteur de structure devient dépendante de la longueur d'onde. Le rapport dispersif est alors défini, pour une paire de longueur d'onde λ_i et λ_j , comme :

$$\frac{\left| {}^{\lambda_i} \tilde{F} \right| - \left| {}^{\lambda_j} \tilde{F} \right|}{\left\langle \left| \tilde{F} \right| \right\rangle} = \frac{\Delta F_{\Delta\lambda}}{F_{moyen}} \quad (25)$$

2.6.3.2.4.2 Acquisition de données MAD

Comme la méthode repose sur la mesure d'intensités de diffraction conduisant à de petites différences d'intensité, elle nécessite des mesures précises.

Pour minimiser les erreurs systématiques sur ces différences, il est préférable d'effectuer toutes les mesures de diffraction sur un seul cristal, de manière à éviter des problèmes de non-isomorphisme et des mises à l'échelle supplémentaires. Les développements récents de la cryocristallographie permettent de généraliser l'utilisation de

cristaux refroidis à 100 K, dans le but de limiter la dégradation du cristal lors des différentes séries de mesures nécessaires. Cette dernière peut constituer une source de non-isomorphisme propre à la méthode MAD.

Le cristal est souvent orienté de manière à mesurer simultanément les intensités des réflexions d'une paire de Bijvoet. Lorsque ce n'est pas possible, la technique de « reverse beam » est utilisée, elle consiste à enregistrer aux différentes longueurs d'onde deux secteurs angulaires séparés de 180° dans des temps rapprochés.

Trois ou quatre ensembles de mesures sont réalisés de manière à optimiser les différences d'intensité dues au signal anomal (Figure II.13).

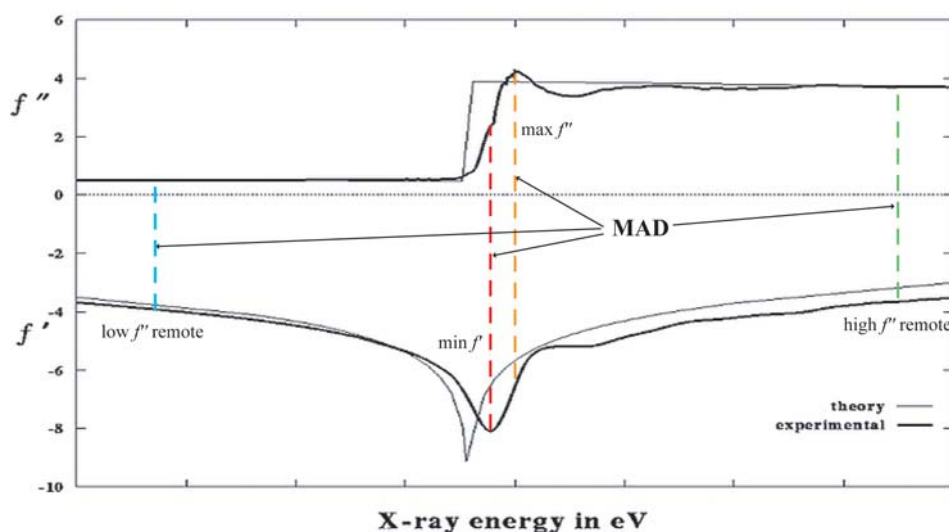


Figure II.13 : Variation des facteurs f' et f'' . Le plus souvent des jeux de données à trois énergies sont mesurés parmi les quatre indiqués sur le dessin : le minimum de f' , le maximum de f'' et une énergie loin du seuil ou en anglais *high f'' remote*.

Les points de mesure incluent la longueur d'onde correspondant au maximum de f'' et la longueur d'onde pour laquelle f' atteint son minimum. Les extremums de f' et f'' ne sont éloignés que de quelques eV. Les autres points de mesure sont éloignés du seuil. Dans le but de se placer au plus près du seuil, les mesures de diffraction sont donc précédées par des mesures de fluorescence X sur le cristal étudié (voir § 2.5.3.2.1.2).

Une grande stabilité en énergie est nécessaire sur les lignes de lumière utilisées pour la mesure d'un jeu de données MAD. En effet, l'information de phase est apportée essentiellement par la mesure des différences de Bijvoet, qui dépendent des variations de f'' . Les contributions de f'' les plus importantes sont obtenues avec les deux séries de mesure au maximum de f'' et au minimum de f' . Contrairement au maximum de la courbe, où une légère instabilité en énergie conduit à une variation de l'amplitude de f'' modérée, la mesure au point d'inflexion, du fait de la pente élevée, est très sensible à une éventuelle instabilité en énergie.

Avec la dégradation du cristal au cours de l'enregistrement, cette sensibilité particulière au degré de précision de l'énergie de mesure constitue une des limites de la méthode.

2.6.3.2.5 Choix des diffuseurs anomaux

Un grand nombre de diffuseurs anomaux peut être utilisé dans le cadre de la méthode MAD. Dans le cadre des acides nucléiques, le brome ($Z=35$) peut être introduit sous la forme de bases bromées.

Lorsque la protéine étudiée contient naturellement un cation métallique (les métalloprotéines comprennent des atomes de Fe, Cu, Ni, Mn, Co, Zn...), cet atome peut être utilisé comme diffuseur anomal.

La substitution des atomes de soufre des méthionines et des cystéines par des atomes de sélénium (Doublé, 1997 ; Strub, 2003) permet d'utiliser le signal anomal au seuil K de cet élément ($\lambda=0,98 \text{ \AA}$) pour obtenir des informations sur la phase. Cette méthode a été proposée par Hendrickson et collaborateurs (Hendrickson et al., 1990). Les protéines contiennent en moyenne une méthionine pour 58 acides aminés. L'amplitude modérée de f'' (une dizaine d'électrons) à ce seuil est compensée par le taux de présence du diffuseur anomal dans la protéine. La méthode s'est généralisée et la majorité des déterminations de structure *de novo* s'effectuent à présent par la méthode MAD sur des protéines sélénées.

Cependant, dans certains cas, si l'expression d'une protéine n'est pas possible chez *E. coli*, l'utilisation de la diffusion anormale nécessite la préparation de dérivés d'atomes lourds sélectionnés en fonction de la longueur d'onde de leur seuil d'absorption et des variations f'' et f' autour de ce seuil. Certains atomes lourds sont utilisés de façon routinière (Hg, Pt, Au, Pb, U etc...) ainsi que certains lanthanides (Yb, Ho, Tb etc...). Ce sont les mêmes atomes que ceux qui sont utilisés dans le cadre des méthodes de remplacement isomorphe (voir § 2.5.3.1.1).

En conclusion, le développement de la méthode MAD et son utilisation en routine ont été rendus possible par :

- le développement des sources de rayonnement synchrotron à travers le monde ; en effet, l'utilisation d'une source de rayonnement réglable à des énergies très précises est la condition *sine qua none* de la réalisation d'une expérience MAD ;
- l'usage généralisé des méthodes de cryoconservation permettant de limiter les dommages causés par le rayonnement auquel les cristaux de protéines sont exposés ;

- et enfin, l'existence d'une méthode de marquage utilisable dans un grand nombre de cas : la substitution des atomes de soufre des méthionines et des cystéines par des atomes de sélénium.

Bien que la méthode MAD, notamment avec l'utilisation des sélénométhionines et peut-être prochainement des sélénocystéines, soit devenue la méthode de prédilection pour déterminer *de novo* la structure de protéines, la méthode SAD joue les outsiders.

2.6.3.3 Cas particulier du SAD

Le problème d'ambiguïté de phase inhérent à la méthode de remplacement isomorphe est la raison d'être des méthodes MIR et MAD. Ainsi la méthode SIR (voir § 2.5.3.1), comportant un seul dérivé lourd, nécessite l'introduction d'une source d'information supplémentaire (deuxième dérivé de la méthode MIR ou diffusion anormale d'un atome lourd présent dans la structure de la méthode SIRAS) pour parvenir à une solution unique. De même, en apparence, la méthode SAD (utilisation de la diffusion anormale des atomes lourds présents dans la structure à une seule longueur d'onde) aboutit à la même ambiguïté sur le choix de la phase. Cette ambiguïté est levée dans la méthode MAD en mesurant d'autres jeux de diffraction à des longueurs d'onde faisant intervenir des comportements anormaux différents de la part des atomes lourds présents (mesures à différentes longueurs d'onde dans un seuil d'absorption).

Dans le cas de la méthode SIR, deux phases possibles sont obtenues pour F_p . Ces deux phases étant équiprobables, une des procédures adoptées consiste à calculer une carte faisant intervenir les deux phases à poids égaux. Une solution F_{best} est ainsi obtenue qui est en fait la superposition de la vraie solution et de la fausse solution pour la phase. La contribution des phases erronées introduit un bruit élevé dans la carte, ce qui rend son interprétation difficile. Cependant, ce bruit a généralement un niveau plus faible que la densité électronique de la protéine. C'est de ce constat que part l'idée de filtrer le bruit autour des régions présentant les densités les plus élevées. Cette « modification » de la densité apporte une information complémentaire qui permet de lever partiellement l'ambiguïté de phase par un processus itératif (ISIR) (Wang, 1985). Cependant, cette méthode repose entièrement sur la qualité des mesures de départ et en particulier sur le bon isomorphisme des données dans le cas du SIR, et sur la précision extrême des mesures dans le cas du SAD.

Dans ce dernier cas, cependant, la situation est légèrement différente. Les deux phases trouvées comme solutions possibles ne sont pas équiprobables. En effet, il existe une

probabilité plus importante pour que la bonne solution pour la phase de la structure totale soit la solution qui est la plus proche de la phase de la sous-structure des atomes lourds. Cette probabilité augmente d'autant que la contribution des atomes lourds au facteur de structure totale est importante. Les phases de la sous-structure des atomes lourds peuvent alors être prises comme phases de départ (Ramachandran et Raman, 1956). La probabilité statistique des phases de la structure de la protéine, en fonction de la contribution des atomes lourds (Sim, 1959), a été utilisé par Hendrickson (Hendrickson et Teeter, 1981) pour résoudre la structure de la crambine à partir de la sous-structure des atomes de soufre, obtenue à la raie $K\alpha$ du cuivre ($\lambda=1,54 \text{ \AA}$). A cette énergie, la partie imaginaire du facteur de diffusion du soufre est de l'ordre de 0,6 électron.

Dans la pratique, les déterminations de structures par la méthode SAD utilisent des méthodes de calculs de probabilités sur les phases (MLPHARE, (Otwinowski, 1991) ; SHARP, (de la Fortelle et Bricogne, 1997)), couplées avec des méthodes de modification de densité (DM, (Cowtan et Main, 1996) ; SOLOMON, (Abrahams et Leslie, 1996) ; RESOLVE, (Terwilliger, 2000).

Cette méthode repose donc entièrement sur la précision des mesures de différences anormales, ainsi que sur leur intensité. Ainsi l'expérimentateur préférera obtenir un jeu de données à une longueur d'onde où la contribution f'' du diffuseur est maximale. La multiplicité des mesures est également un facteur essentiel pour la réussite de la méthode (Dauter et al., 1999).

Le développement de la méthode SAD fait partie intégrante d'une démarche générale, qui vise à optimiser le temps pris pour la détermination de structure de protéine *de novo*. En effet, l'ère de la génomique structurale incite à trouver des moyens rapides et applicables dans la plupart des cas pour déterminer de nouvelles structures. Le soufre est l'atome le plus lourd naturellement présent dans l'ensemble des acides aminés. Une stratégie, différente de celle d'Hendrickson (Hendrickson et Teeter, 1981), consiste à augmenter la longueur d'onde pour se rapprocher du seuil K du soufre ($\lambda=5,02 \text{ \AA}$), afin d'augmenter sa contribution anormale. C'est une des voies prometteuses de l'utilisation des grandes longueurs.

2.6.4 Méthodes directes

Ces dernières années, les méthodes depuis longtemps utilisées avec succès pour la détermination des structures de petites molécules commencent à être appliquées aux protéines. Ces méthodes se sont montrées efficaces pour de petites protéines à condition

d'avoir des données de diffraction à très haute résolution (autour de 1,2 Å). Elles sont fondées soit sur le principe des méthodes directes (Sheldrick, 1990 ; Hauptman, 1997), soit sur celui de la déconvolution des fonctions de Patterson (Sheldrick, 1997). Malheureusement, les conditions d'utilisation de ces méthodes sont très restrictives. Elles sont néanmoins très utiles lorsqu'il s'agit de trouver la position d'atomes lourds dans la maille à partir de données de différences isomorphes ou anormales.

2.7 Modification de densité

Lors de la détermination des phases par une méthode de phasage expérimentale (MIR ou MAD), aucune hypothèse n'a encore été faite quant au contenu de la maille. Ces phases ne sont donc pas biaisées par un quelconque modèle. Par contre, elles ne sont pas exactes, du fait d'un certain nombre d'erreurs expérimentales. De plus, il se peut que les phases aient été calculées à une certaine résolution et qu'un jeu de données natif puisse être à plus haute résolution. Les premières cartes de densité électronique calculées à partir de ces phases peuvent ne pas être du tout interprétables. Il est alors possible ou souhaitable d'introduire quelques contraintes générales sur la densité électronique, de manière à améliorer les phases expérimentales. Ce sont les étapes de modification de densité électronique qui regroupent principalement les procédures suivantes :

- l'aplatissement de solvant ;
- l'*histogram matching*
- la moyennation dans le cas où il existe des symétries non cristallographiques.

Toutes ces méthodes permettent d'améliorer les phases expérimentales car elles introduisent des relations entre les différents facteurs de structure, et donc des contraintes sur les phases.

2.7.1 Principe de l'aplatissement de solvant

Cette procédure est basée sur le fait que le solvant, non ordonné dans le cristal, ne participe pas à la diffraction, sauf à basse résolution. La densité électronique correspondant aux zones de solvant doit donc être plate, ce qui n'est pas le cas, du fait des erreurs de phase qui introduisent des fluctuations. Or, la densité électronique moyenne correspondant aux zones de protéine est plus élevée que celle correspondant au solvant et les fluctuations dans les zones de solvant seront plus faibles, bien qu'elles puissent bruite une carte au point de la rendre non interprétable à l'œil. La méthode décrite par Wang (1985) consiste à discriminer les zones de solvant des zones de protéine, sur la base de la valeur locale moyenne de la

densité électronique, en tenant compte du pourcentage de solvant du cristal pour définir le seuil de coupure. Cela permet de déterminer un masque ayant pour valeur 1 pour les zones de protéine et 0 pour les zones de solvant. Il suffit alors d'appliquer ce masque, puis de recalculer des phases à partir de cette nouvelle carte de densité électronique nivelée. Ensuite, il est possible de refaire quelques cycles de détermination de masque et nivellement jusqu'à convergence du facteur R_{libre} (Figure II.14). Ce facteur R_{libre} est le même que celui utilisé pour suivre l'affinement (voir § 2.7.3). Il mesure l'accord entre des F_{obs} mis de côté et n'intervenant pas dans les calculs de carte de densité électronique et les F_{calc} restaurés par la procédure qui introduit des relations entre les différents facteurs de structure et donc des contraintes sur le modèle comme sur les phases. Cette méthode sera d'autant plus puissante que le pourcentage de solvant dans le cristal est élevé.

2.7.2 Symétries non cristallographiques – moyennation

2.7.2.1 Symétries non cristallographiques

Il est possible que l'unité asymétrique de la maille cristalline contienne plusieurs motifs identiques. Il s'agit donc de symétries non cristallographiques qui ne sont valables que localement, aux alentours des motifs concernés (voir § 2.4.2).

2.7.2.2 Moyennation

L'existence de symétries non cristallographiques permet de remplacer la densité électronique de chaque motif par la densité électronique moyenne correspondante. Cette procédure est extrêmement puissante pour introduire des contraintes dans les facteurs de structure calculés, et donc améliorer drastiquement les phases, ou même procéder à une extension de phases sur des facteurs de structure à plus haute résolution, pour lesquels aucune phase expérimentale n'a pu être déterminée. Cette méthode nécessite la détermination préalable des opérateurs de symétrie non cristallographique reliant les différents motifs de l'unité asymétrique, ainsi que le masque décrivant la zone correspondant au motif concerné par la symétrie non cristallographique.

2.7.3 Procédure de modification de densité électronique

L'étape de modification de densité est une procédure itérative (voir Figure II.14). A chaque cycle, de nouvelles phases calculées sont obtenues et combinées avec les phases expérimentales de façon à éviter le biais que peut introduire cette procédure. Cela est effectué jusqu'à convergence du facteur R_{libre} . Normalement, il est possible d'obtenir une carte de

densité électronique de bien meilleure qualité que la carte expérimentale du départ, surtout s'il a été possible de réaliser une moyennation sur plusieurs molécules de l'unité asymétrique.

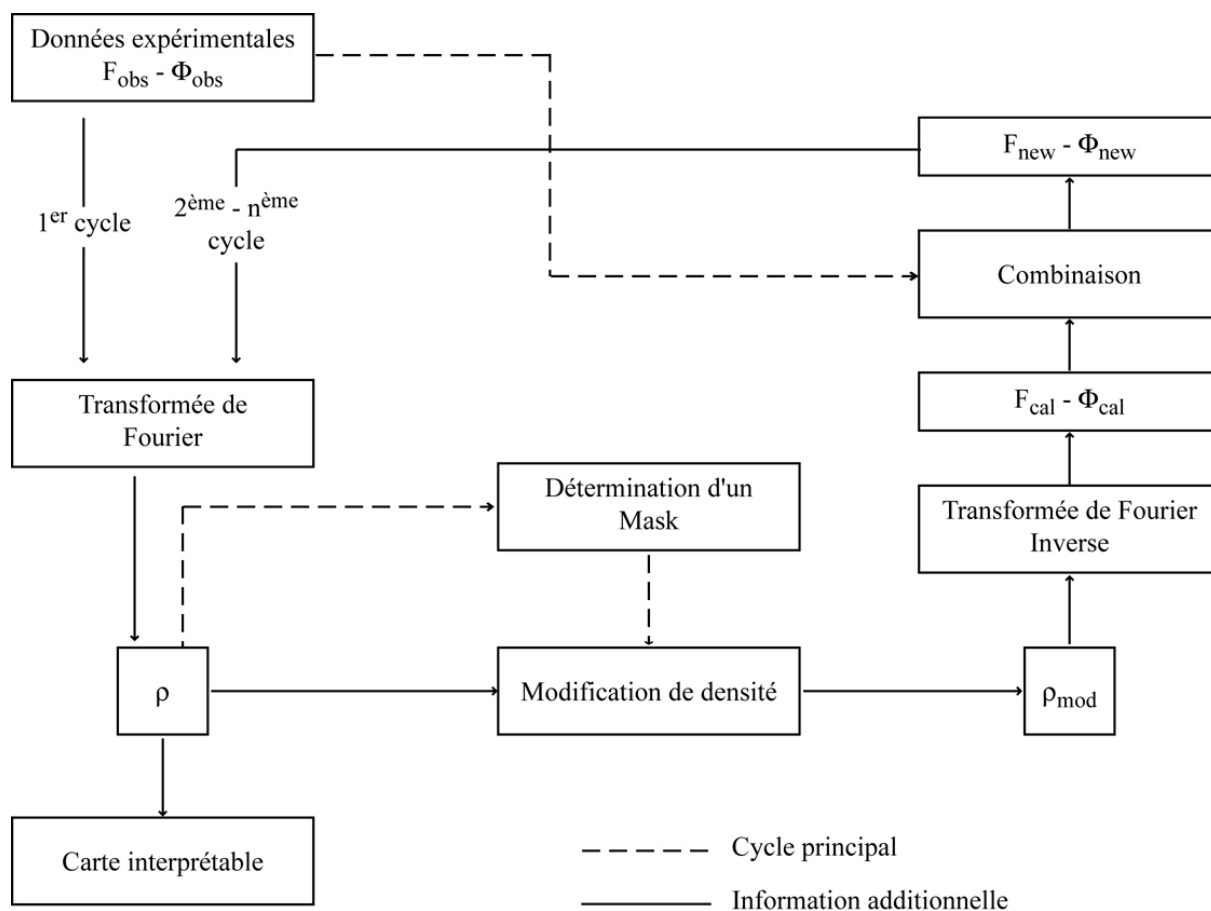


Figure II.14 : Schéma de la procédure de modification de densité électronique.

2.8 Affinement et construction de la structure

Par l'étape de détermination de la structure par remplacement moléculaire ou par modélisation dans une carte de densité électronique interprétable, nous disposons d'un modèle approximatif de la structure que nous cherchons à déterminer. Les facteurs de structure calculés à partir de ce modèle sont généralement plus ou moins en accord avec les facteurs de structures observés par les expériences de diffraction des rayons X.

Il est donc nécessaire d'améliorer cet accord par des techniques de reconstruction manuelle avec les programmes de visualisation et de manipulation des modèles atomiques et des cartes de densité électronique et par des techniques d'affinements numériques des paramètres du modèle (position et agitation de chaque atome). L'alternance de la reconstruction manuelle et de l'affinement numérique permet : 1) de modifier le modèle lorsque celui-ci est trop éloigné de la structure recherchée et donc quand les algorithmes

d'affinement ne sont pas capables de converger seuls ; 2) de compenser par l'affinement numérique les erreurs induites par la reconstruction manuelle.

Pour affiner l'ensemble des nombreux paramètres de position et d'agitation de la structure atomique d'une protéine, le nombre d'observations doit être bien supérieur au nombre de ces paramètres. Pour les résolutions couramment rencontrées en cristallographie des protéines, nous sommes confrontés au problème du nombre insuffisant d'observations (les facteurs de structures) par rapport au grand nombre de paramètres structuraux à affiner. Ce problème de surdétermination est plus ou moins ressenti suivant la résolution maximale des données observées. En effet, pour chaque méthode d'affinement, un rayon de convergence peut être défini : si la position initiale d'un atome est en dehors de cette zone de convergence, le système d'affinement ne sera pas capable de l'amener à sa vraie position. Ce rayon de convergence dépend de la résolution des données et donc du taux de surdétermination. En général, l'affinement numérique converge plus facilement avec des données ayant une résolution maximale proche de l'Angström que celui avec des données ayant une résolution de l'ordre de 3 Å pour lesquelles il n'y a qu'une observation par paramètre affiné. La faible surdétermination du système d'affinement peut être compensé par l'utilisation de nouvelles observations telles que les valeurs des distances et des angles de liaisons atomiques déterminées avec précision pour les petites molécules (Engh et Huber, 1991) ou encore l'identité par les symétries non-cristallographiques. Ces observations supplémentaires serviront de contraintes stéréochimiques ou cristallographiques.

Il existe deux méthodes d'affinement très répandues : les méthodes de minimisation d'un résiduel (moindres carrés ou énergie) ou de dynamique moléculaire.

2.8.1 L'affinement par minimisation

2.8.1.1 Le terme à minimiser

Cette méthode d'affinement permet de minimiser un terme qui tient compte des différences entre les valeurs des différents facteurs calculées à partir du modèle (facteurs de structures et valeurs stéréochimiques) et les valeurs observées de ces mêmes facteurs (facteurs de structures mesurés, contraintes stéréochimiques).

La fonction à minimiser s'exprime par :

$$S = W_{\text{xray}} F_{\text{xray}} + \sum_i W_i \sum_j \frac{1}{\sigma_j^2} (P_{\text{idj}} - P_{\text{obsj}})^2 \quad (26)$$

où P_{obsj} sont les facteurs structuraux observés et P_{idj} leur valeur calculée à partir du modèle. Ils peuvent être de différentes natures comme les contraintes stéréochimiques (longueurs, angles des liaisons interatomiques, répulsions de Van der Waals entre atomes non liés), cristallographique (similarité par symétrie non cristallographique) ou encore contraintes sur la variation lors de l'affinement des paramètres de position ou d'agitation.

Les facteurs W_i et σ_j permettent de contrôler la pondération de chaque terme. Ainsi il est possible d'éliminer ou de renforcer une catégorie de paramètres, en particulier en régularisant la géométrie d'un modèle en lui imposant W_{xray} égal à 0.

Le terme F_{xray} qui rend compte du désaccord entre les facteurs de structure calculés et observés peut s'exprimer de deux façons différentes selon qu'il utilise les procédures de minimisation par moindres de carrés ou de maximisation de vraisemblance.

2.8.1.2 Les moindres carrés

Dans le cas de la minimisation par moindres carrés, le terme des rayons X s'écrit :

$$F_{\text{xray}} = \sum_{\text{hkl}} \frac{1}{\sigma_F^2} (|F_{\text{obs}}| - |F_{\text{calc}}|)^2 \quad (27)$$

L'avantage de cette méthode d'optimisation des paramètres réside dans la vitesse de calcul des équations de l'algorithme et d'une convergence rapide. Toutefois, si la diminution du résiduel résulte de l'amélioration du modèle, il compense aussi les erreurs des données mesurées par des erreurs dans le modèle affiné.

2.8.1.3 Le maximum de vraisemblance

La méthode de maximisation de vraisemblance part d'un principe différent. La vraisemblance du modèle ou des facteurs de structure calculés est définie comme étant la probabilité de faire la mesure des F_{obs} si l'hypothèse du modèle est juste :

$$\prod_i P(F_{\text{obsi}} / \text{tous les } F_{\text{calc}}) \quad (28).$$

La méthode de maximum de vraisemblance propose alors de minimiser le terme : $F_{\text{xray}} = -\sum_i \log P(F_{\text{obsi}} / \text{tous les } F_{\text{calc}})$ (27) et en estimant cette fonction de probabilité à partir du jeu de réflexions utilisé pour le calcul du R_{libre} (voir § II.2.7.4) afin d'éviter un effet de biais induit par le modèle affiné.

Comparée à la méthode des moindres carrés, la maximisation de la vraisemblance permet d'atteindre plus facilement le minimum des facteurs R et R_{libre} . Toutefois, les affinements par cette méthode ont montré une capacité à améliorer un modèle basé sur des données à haute résolution plutôt qu'à basse résolution à cause du faible nombre de réflexions (Murshudov et al., 1997)

2.8.1.4 Le terme d'énergie

Certains algorithmes d'optimisation du modèle utilisent une fonction assimilée à l'énergie potentielle qui est l'énergie potentiel de la molécule complétée par le terme de rayons X. Cette fonction est très analogue au résiduel des moindres carrés :

$$S = W_{\text{xray}} \sum_{hkl} (|F_{\text{obs}}| - |F_{\text{calc}}|)^2 + \sum_i W_i \sum_j \frac{1}{2} K_j (P_{\text{idj}} - P_{\text{obsj}})^2 \quad (29)$$

Les facteurs K_j représentent explicitement des constantes de force et utilisent les modèles physiques de l'oscillateur harmonique ou encore des interactions ou répulsions des atomes (potentiels de Lennard-Jones ou coulombiens).

2.8.2 La dynamique moléculaire

Dans le cas où la structure du modèle est très éloignée de la vraie structure, la fonction d'énergie peut être bloquée dans un minimum local au lieu d'atteindre le vrai minimum. Pour éviter ce genre de désagrément, les méthodes d'affinement par dynamique moléculaire peuvent être utilisées. Le principe est de simuler le comportement dynamique des atomes par les lois de la mécanique newtonienne. A chaque atome de la molécule est assigné une température élevée, donc une vitesse, ce qui confère à la molécule une certaine énergie cinétique. Le modèle aura ainsi une énergie suffisante pour franchir des barrières d'énergie potentielle importantes. La diminution lente de la température, ou recuit simulé, amène ensuite le système vers des minima d'énergie différents de son point de départ.

Cette méthode permet, de par son large rayon de convergence, de corriger des erreurs importantes dans le modèle de départ. Néanmoins, elle demande des temps de calcul très longs et il est souvent plus rapide d'alterner reconstruction manuelle et affinement par minimisation pour corriger les erreurs structurales qui pourraient l'être par cette méthode de recuit simulé.

2.8.3 Les programmes utilisés

Pour l'affinement des différentes structures, nous avons employé deux des programmes les plus couramment utilisés : REFMAC (*Macromolecular Refinement program*) développé par (Murshudov et al., 1997) et CNS (*Crystallography and NMR System*) développé par (Brunger et al., 1998).

2.8.3.1 REFMAC

Ce programme disponible avec la suite de programmes CCP4 (CCP4, 1994) a spécialement été développé pour affiner les structures par les méthodes de minimisation.

Ce programme suit une procédure tout à fait classique :

Tout d'abord, à partir des coordonnées atomiques du modèle, il détermine les contraintes géométriques (liaisons atomiques...), calcule les facteurs de structure, en module et en phase, et les met à l'échelle par rapport aux facteurs de structure observés. Puis vient proprement dite l'étape d'affinement qui peut se faire soit par la méthode des moindres carrés ou de maximum de vraisemblance. REFMAC minimise la fonction choisie par la méthode du gradient conjugué : dans l'espace des paramètres affinés, le minimum de la fonction est atteint en faisant varier les valeurs des paramètres afin de se déplacer dans la direction opposée au gradient de cette fonction. Plusieurs cycles sont nécessaires pour atteindre un minimum local. Grâce à différentes options, le programme propose soit d'affiner la structure en corps rigides (modèle divisé en morceaux considérés comme rigides), ou d'affiner l'ensemble des coordonnées avec ou sans le facteur d'agitation de chaque atome.

Le programme, à la fin de cette procédure, calcule les nouvelles coordonnées du modèle affiné et ses facteurs de structures associés.

L'intérêt principal de REFMAC est l'algorithme de minimisation par maximum de vraisemblance développé par G. Murshudov.

2.8.3.2 CNS

Ce programme développé par A. Brunger sur la base d'XPLOR (Brunger, 1992b) utilise la minimisation de l'énergie globale du modèle.

Outre les méthodes de minimisation, il propose une méthode d'affinement basée sur la dynamique moléculaire (recuit simulé).

Le grand avantage de ce programme est sa convivialité. En effet, A. Brunger a développé une interface HTML qui permet de mieux appréhender les différents paramètres de l'affinement. Il propose aussi tout un ensemble de scripts différents pour affiner une structure moléculaire

(affinement des coordonnées atomiques, des facteurs d'agitation individuellement ou groupés, dynamique...). Par rapport à X-PLOR, de nouvelles fonctionnalités ont aussi été ajoutées. Il est possible maintenant avec CNS de résoudre une structure par le phasage avec les atomes lourds, d'améliorer les phases expérimentales par les méthodes de modification de densités ou encore d'affiner les structures RMN.

2.8.4 Cartes de densité électronique

Pour la visualisation de la densité électronique, il n'est pas souhaitable d'utiliser directement le résultat de la transformée de Fourier calculé à partir des coefficients

$|F_{\text{obs}}| \cdot e^{i2\pi\phi_{\text{calc}}}$. En effet, l'information structurale est essentiellement contenue dans le terme de phase (voir (2) et (3)). Les phases calculées à partir du modèle biaisent donc fortement le résultat de la transformée (l'image de la densité électronique) par l'inexactitude de celui-ci.

Pour rendre compte des différences entre le modèle et l'image de la densité électronique, il est préférable d'inclure la différence entre les modules des facteurs de structure observés et calculés dans le calcul de la transformée de Fourier. Deux types de cartes de densité électronique, dites "cartes différences", sont ainsi calculés pour la visualisation et la reconstruction manuelle. Ils utilisent les coefficients :

$$(2|F_{\text{obs}}| - |F_{\text{calc}}|)e^{i2\pi\phi_{\text{calc}}} \text{ et } (|F_{\text{obs}}| - |F_{\text{calc}}|)e^{i2\pi\phi_{\text{calc}}} \quad (30)$$

La carte $2F_{\text{obs}} - F_{\text{calc}}$ peut être considérée comme la densité électronique de la molécule à résoudre. Elle est a priori moins biaisée par l'erreur du modèle grâce à l'information apportée par la différence $F_{\text{obs}} - F_{\text{calc}}$.

La carte $F_{\text{obs}} - F_{\text{calc}}$ est la carte de densité résiduelle qui permet de mettre en évidence les différences entre la densité du modèle et celle observée. Elle permet aussi d'identifier des erreurs grossières sur la position des atomes ou encore la présence d'atomes manquant dans le modèle.

Par les analyses statistiques de phases calculées, il est possible de calculer différents coefficients qui estiment l'exactitude de celles-ci et qui seront inclus dans le calcul des cartes de densité électronique ($2m \cdot F_{\text{obs}} - D \cdot F_{\text{calc}}$ et $mF_{\text{obs}} - F_{\text{calc}}$). Le coefficient m , qui est une estimation de la précision de la phase, et D sont des poids appliqués aux termes de phase et aux modules des facteurs de structure calculés. Ils prennent en compte les erreurs apportées par l'incomplétude du modèle et l'inexactitude des positions des atomes et permet d'améliorer la qualité des cartes différences.

Pour visualiser et manipuler ces cartes de densités électroniques, ainsi que les modèles moléculaires, nous avons essentiellement utilisé le logiciel TURBO (Roussel et al., 1989).

2.8.5 Critères d'évaluation du modèle

Il existe plusieurs critères pour évaluer la qualité d'un modèle affiné :

- **Le facteur R** peut donner une idée de la qualité du modèle en mesurant l'accord entre les facteurs de structure observés et les facteurs de structure calculés :

$$R = \frac{\sum_{hkl} ||F_{obs}| - |F_{calc}||}{\sum_{hkl} |F_{obs}|} \quad (31)$$

Toutefois, les méthodes d'affinement tentent de minimiser au maximum ce facteur et il peut arriver que les erreurs du modèle compensent les erreurs des mesures lorsque le système observables-mesures est peu surdéterminé.

Une estimation moins biaisée de la qualité du modèle peut être définie par le facteur R_{libre} calculé de la même façon que le facteur R mais sur un ensemble de réflexions choisies aléatoirement (5 à 10 %) et non incluses dans l'affinement (Brunger, 1992a). Ce facteur, fortement corrélé à la précision des phases, permet de détecter de multiples types d'erreurs dans la structure. Il peut aussi être utilisé pour la détermination de certains paramètres de l'affinement comme l'optimisation des poids utilisés. L'écart entre les facteurs R et R_{libre} est finalement un moyen de qualifier l'état de surdétermination du système et donc de la résolution maximale des données : plus le rapport entre le nombre d'observations et le nombre de paramètres à affiner est faible, plus cet écart est important (Tickle et al., 1998). Ainsi dans le cas d'un suraffinement du modèle lorsque le nombre d'observations n'est pas suffisant, le R_{libre} aura tendance à augmenter pendant que le facteur R diminue (Kleywegt and Brunger, 1996).

- **Les déviations moyennes** des longueurs et angles de liaisons entre atomes du modèle estiment la qualité stéréochimique de celui-ci. La qualité stéréochimique d'une structure est considérée comme bonne quand ses déviations ne sont pas supérieures aux variations physiques des longueurs et des angles de liaison déjà observées pour les petites molécules (environ 0.02 Å pour les longueurs et 2 ° pour les angles) (Engh and Huber, 1991).

- **Le diagramme de Ramachandran** permet lui aussi de juger la qualité stéréochimique du modèle affiné (Ramakrishnan et Ramachandran, 1965). En raison de l'encombrement stérique des chaînes latérales des aminoacides, les liaisons peptidiques ne

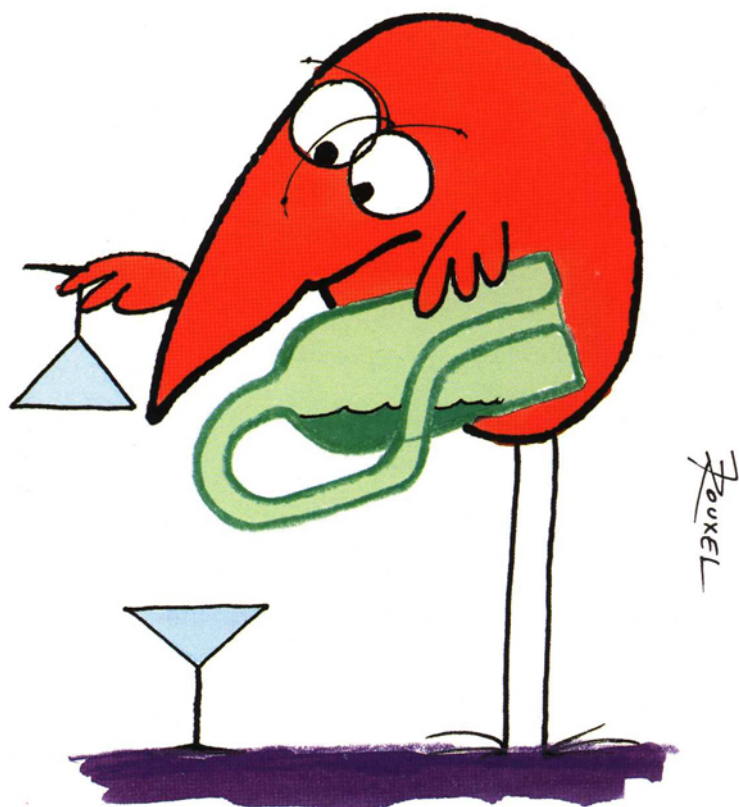
peuvent adopter qu'un certain nombre limité de conformations. Pour chaque résidu, ces conformations sont caractérisées par les angles Φ et Ψ de la liaison peptidique et sont représentées sur un graphe : le diagramme de Ramachandran qui présente les régions énergétiquement favorables, les régions permises et celles non autorisées.

Les chaînes latérales sont elles aussi soumises à des conformations énergétiquement favorables qui seront fortement représentées parmi les différentes conformations rencontrées dans les structures de protéines. Le même type de diagramme permet de valider leur stéréochimie.

- Enfin, **les contacts atomiques, la longueur et l'angle des liaisons hydrogènes potentiels** permettent lorsque ceux-ci ne sont pas corrects d'identifier les régions erronées du modèle.

Résultats

Les devises Shadok



*S'IL N'Y A PAS DE SOLUTION
C'EST QU'IL N'Y A PAS DE PROBLÈME.*

Chapitre III : Résultats

1 Les premières données

1.1 Cristallisation

La protéine IRP1 humaine a été surexprimée dans *E. coli* et purifiée selon un protocole établi par nos collaborateurs (Brazzolotto et al., 1999). Plusieurs échantillons nous ont été fournis pour nos expériences de cristallogenèse. La présence de l'agrégat Fe-S rend la protéine instable à l'air et pour cette raison toutes les expériences de cristallogenèse ont été réalisées en milieu anaérobie. Les premiers essais de cristallisation ont été basés sur les conditions qui avaient permis la cristallisation de l'aconitase mitochondriale de porc (Robbins et al., 1982). Yvain Nicolet, qui est à l'origine du projet au sein du LCCP, a fourni un très gros travail de dégrossissement en effectuant plusieurs gammes en sulfate d'ammonium à différents pH grâce à la technique de la diffusion de vapeur par goutte suspendue (voir § II.1.2). Le problème a été que seuls des oursins étaient obtenus et tous les essais d'ensemencement ainsi que tous les cocktails d'additifs n'ont rien changé à cet état de fait (Figure III.1a). L'analyse, par électrophorèse, des échantillons a révélé un manque d'homogénéité dans les préparations.

Claudine Darnault a pris la suite du projet avec de nouveaux lots de purification et tout en essayant de reproduire les résultats obtenus pour l'aconitase mitochondriale, elle a utilisé l'échantillonnage en matrice dispersée proposée par Jancarick et Kim (1991). Pour ces expériences la protéine a été concentrée à 20 mg/ml dans un tampon contenant 10 mM Hepes pH 7,4 et incubée pendant deux heures à 293 K avec 200 μ M de citrate tri-sodium. Les gouttes suspendues ont été réalisées avec 1 μ l de solution de protéine auquel 1 μ l du réservoir a été ajouté, le tout étant placé sur 1 ml de solution réservoir. La condition 41 du *Crystal Screen I* (Hampton Research), c'est-à-dire 100 mM Hepes 7,5, 10% isopropanol et 20% w/v PEG 4000, a permis d'obtenir des plaquettes mélangées à des aiguilles, le tout étant agrégé (Figure III.1b). Les cristaux ont été améliorés en faisant varier le pH et la concentration des deux agents précipitants. Des agrégats de grandes plaquettes ont ainsi pu être obtenus dans une solution de 100 mM Tris-HCl pH 8,1, 10% isopropanol, 18% PEG 4000 et 2 mM Na-dithionite (Figure III.1c). A partir de ces agrégats deux plaquettes monocristallines ont été séparées et congelées directement dans la boîte à gants dans un tampon légèrement plus concentré en agents précipitants que celui ayant permis la cristallisation auquel nous avons

ajouté du glycérol comme cryoprotectant (100 mM Tris-HCl pH 8,1, 20% isopropanol, 20% PEG 4000, 20% glycérol et 2 mM Na-dithionite). Les plaquettes, ainsi récupérées, avaient une taille proche des 250 μm par 150 μm par 20 μm et nous leur avons attribué les noms arbitraires de cristal 16 (X16) et de cristal 40 (X40).

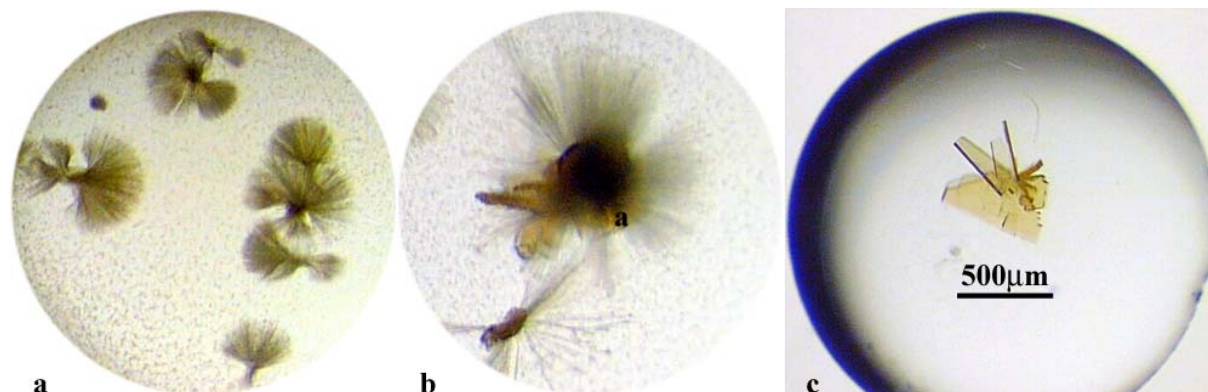


Figure III.1 : Evolution de la morphologie des cristaux. a) représente les cristaux en forme d'oursin obtenus dans les conditions de cristallisation de l'aconitase mitochondriale de porc, b) montre le résultat de la solution 41 du *crystal screen I* (Hampton Research) et c) l'optimisation de cette même condition.

1.2 Cristaux natifs

1.2.1 Collecte des données et statistiques

Les premiers jeux de données natifs ont été enregistrés à l'ESRF sur la ligne de lumière ID14-4 et l'ensemble a été traité à l'aide du programme XDS (Kabsch, 1993). Les tableaux suivant (Tableau III.1 et III.2) regroupent les principales statistiques après mise à l'échelle.

Cristal	Grp	a (Å)	b(Å)	c(Å)	$\alpha(^{\circ})$	$\beta(^{\circ})$	$\gamma(^{\circ})$
16	P2 ₁	63,9	215,5	64,0	90	72	90
40	C2	103,5	75,1	214,7	90	90	90

Tableau III.1 : Paramètres de maille des cristaux natifs.

Les cristaux 16 et 40 sont tous les deux monocliniques mais ils n'ont pas le même groupe d'espace, ce qui peut être assez surprenant car ils proviennent tous les deux d'un même agrégat de cristaux.

Cristal	16	40
Ligne de lumière ESRF	ID14-4	ID14-4
Résolution	20-1,85	20-2,0
Nombre d'observations	547170	355541
Réflexions uniques	270649	187067
Complétude (%)	97,5 (92,1)	88,6 (56,8)
Redondance	2,0 (1,6)	1,9 (1,4)
$\langle I/I \rangle$	8,6 (3,0)	8,5 (4,0)
R_{sym} (%)	6,0 (26,2)	6,1 (14,5)
R_{meas} (%)	8,0 (35,0)	8,4 (20,1)

Tableau III.2 : Statistiques des données natives indexées et intégrées par XDS (Kabsch, 1993). Les valeurs entre parenthèses correspondent à la dernière coquille de résolution.

$$R_{\text{sym}} = R_{\text{merge}} = \sum_h \left| \hat{I}_h - I_{h,i} \right| / \sum_h \sum_i^{n_h} I_{h,i}, \quad R_{\text{meas}} = \sum_h \sqrt{\frac{n_h}{n_h - 1}} \sum_i^{n_h} \left| \hat{I}_h - I_{h,i} \right| / \sum_h \sum_i^{n_h} I_{h,i} \quad \text{avec} \quad \hat{I}_h = \frac{1}{n_h} \sum_i^{n_h} I_{h,i}$$

(Diederichs et Karplus, 1997).

1.2.2 Nombre de molécules par unité asymétrique

Le groupe d'espace $P2_1$ possède deux positions équivalentes alors que le groupe d'espace $C2$ en possède quatre. Généralement, le volume de Matthews (V_m) est compris entre 2 et 4 Å³/Da (Matthews, 1968). Dans notre cas les volumes des mailles sont respectivement de 838174 Å³ pour le $P2_1$, de 1668830 Å³ pour le $C2$ et le poids moléculaire de la protéine est de 98400 Da. Le tableau suivant (Tableau III.3) présente le V_m correspondant pour différents nombres de molécules par unité asymétrique.

Nb mol /asym	X16 V_m (Å ³ /Da)	X40 V_m (Å ³ /Da)
1	4,3	4,2
2	2,1	2,1
3	1,4	1,4

Tableau III.3 : Volume de Matthews correspondant aux deux jeux natifs.

Il est possible de dire, sur la base de ce calcul, que l'unité asymétrique, des deux formes cristallines, contient très vraisemblablement 2 molécules d'IRP1.

1.3 Enregistrement de jeux MAD

A ce stade, il paraît bon de rappeler que la faible homologie de séquence avec l'aconitase mitochondriale de bœuf ne nous permettait pas d'utiliser la technique du remplacement moléculaire pour résoudre le problème de la phase.

Suite à ces deux premiers jeux de bonne qualité, nous avons décidé d'essayer de résoudre le problème de la phase en enregistrant les jeux nécessaires à une expérience de MAD. Afin de limiter les problèmes d'isomorphisme nous nous sommes focalisés, dans un premier temps, sur le cristal 40. Comme l'ensemble des enregistrements a été effectué sur FIP-BM30a à l'ESRF (voir § II.2.2.1), nous avons pu utiliser le détecteur de fluorescence résolu en temps (Röntec, Germany) qui est présent sur la ligne et qui permet, en fonction de la longueur d'onde d'excitation, de vérifier la présence de diffuseurs anomaux dans le cristal. Cette mesure, rapide, permet de pouvoir cribler un grand nombre d'échantillons avant de choisir celui qui semble avoir le meilleur signal de fluorescence (Figure III.2).

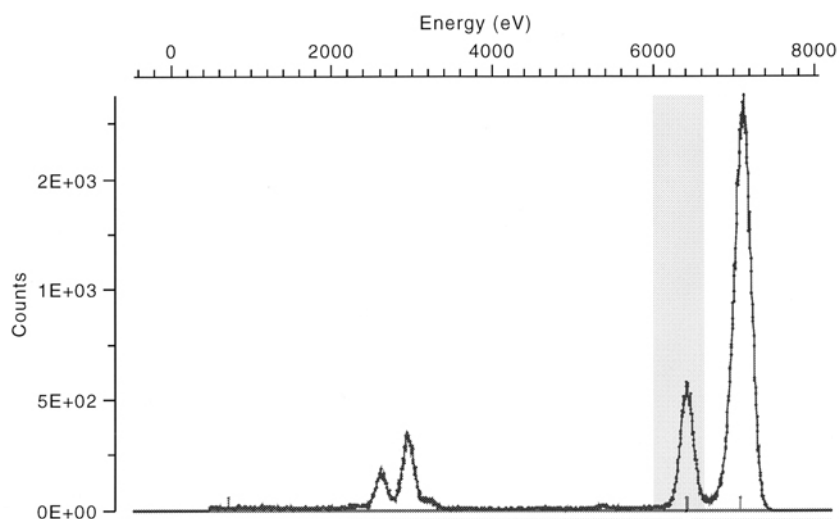


Figure III.2 : Spectre de fluorescence au seuil d'excitation du fer. Le dernier pic correspond au signal diffusé alors que celui qui est grisé montre la raie de fluorescence $K\alpha$ du fer dont l'énergie est la signature de la présence de fer.

Le choix des longueurs d'ondes a nécessité la détermination de la variation des f' et f'' en fonction de la longueur d'onde au voisinage du seuil d'absorption du diffuseur anomal. Cette détermination est réalisée à partir du spectre de fluorescence X du cristal en fonction de l'énergie des photons X. La mesure de fluorescence permet de déterminer directement la variation du facteur f'' . Les valeurs de f' peuvent alors être calculées par la relation de Kramers-Kronig (Kronig et al., 1928) (voir équation 17). Nous avons donc choisi les longueurs d'ondes qui correspondaient au maximum de f'' , afin d'avoir le plus d'électrons

anomaux possible, au minimum de f'' et une dernière collecte a été réalisée à haute énergie pour la *remote* à 0,97 Å (Figure III.3).

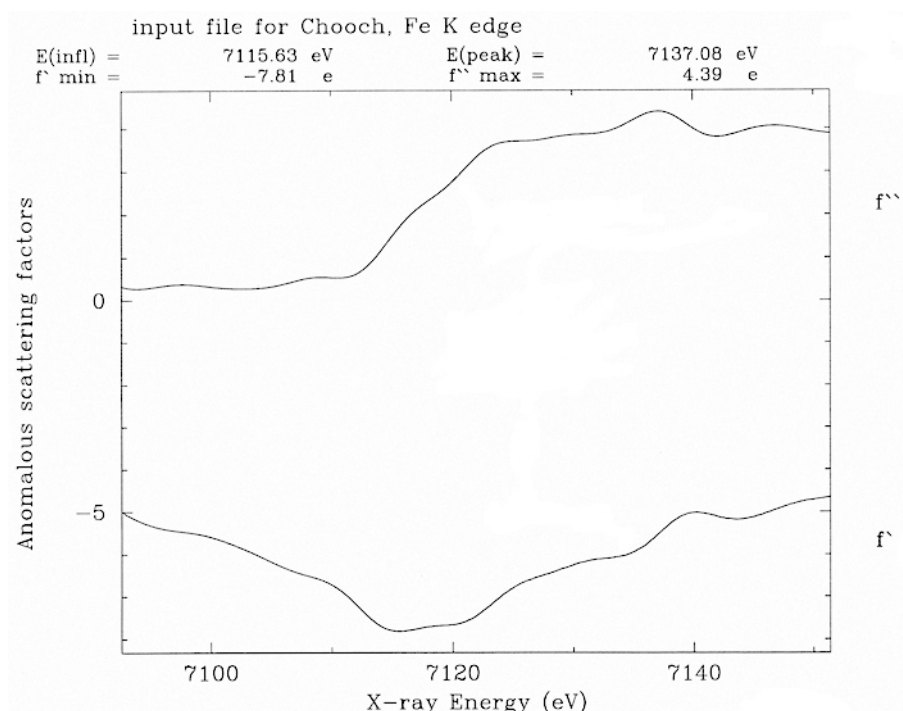


Figure III.3 : Spectre de fluorescence au seuil K du fer mesuré sur le cristal afin de tenir compte de l'influence des ligands. Trois longueurs d'ondes ont été utilisées pour les enregistrements, 0,972 Å, 1,738 Å et 1,743.

L'indexation des données par le programme XDS (Kabsch, 1993) a permis de déterminer la maille cristalline : $a=103,49$ Å, $b=75,08$ Å, $c=214,68$ Å et $\alpha=\beta=\gamma=90^\circ$ dans le groupe d'espace C2. Les statistiques sont reprises dans le tableau ci-dessous.

Cristal	Remote	max f''	min f'
Ligne de lumière ESRF	FIP-BM30a	FIP-BM30a	FIP-BM30a
Longueur d'onde (Å)	0,972	1,738	1,743
f'/f''	-/-	-4,5/3,9	-6,1/3,9
Résolution	30-2,0	25-2,9	50-3,0
Nombre d'observations	494359	203604	229393
Réflexions uniques	102640	68470	63288
Complétude (%)	92,2 (61,3)	95,7 (84,8)	98,5 (92,7)
Redondance	4,8 (2,1)	3,0 (2,4)	3,6 (3,0)
$\langle I/I \rangle$	13,5 (5,19)	12,34 (4,45)	14,08 (5,47)
R_{sym} (%)	8,2 (15,7)	7,5 (21,8)	7,6 (19,4)
R_{meas} (%)	9,0 (19,9)	9,1 (27,1)	9,0 (23,4)

Tableau III.4 : Statistiques des données MAD après leur mise à l'échelle.

1.4 Détermination de la position des agrégats dans l'unité asymétrique

Les données des différents jeux MAD ont ensuite été traitées avec la suite de programme de CCP4 (CCP4, 1994). Etant donné le contenu métallique de la protéine IRP1 et le contenu de l'unité asymétrique, la détermination de la position des atomes de fer individuels revient à trouver $2 \times 4 = 8$ sites indépendants. Comme les atomes de fer sont regroupés en deux agrégats $[4\text{Fe}-4\text{S}]$, nous avons décidé de les considérer comme des super atomes dans un premier temps. Cette approximation est valable jusqu'à une résolution de 5 Å, qui correspond à la taille des agrégats (Chabrière et al., 1999). Le nombre de positions indépendantes à déterminer se trouvait donc réduit à deux. La carte de Patterson nous a néanmoins réservé quelques surprises (Figure III.4).

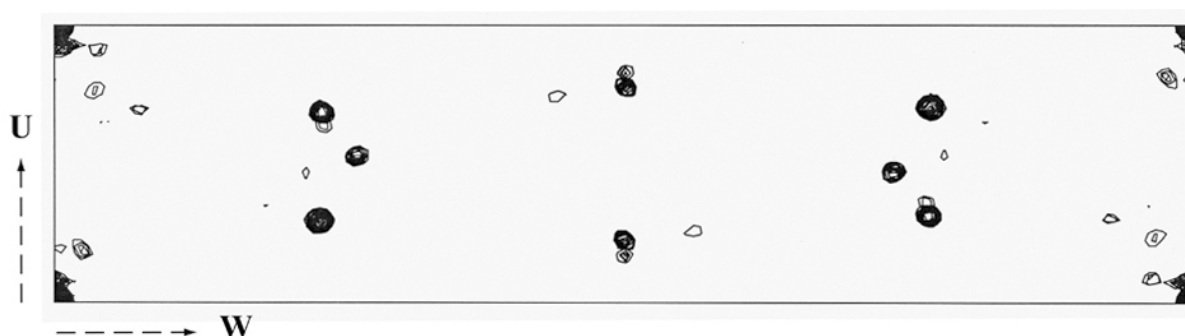


Figure III.4 : Carte de Patterson du cristal 40 à 3,5 Å, section de Harker $v=1/2$.

La carte de Patterson en C2 du cristal 40 s'explique très bien avec trois sites, ce qui n'est pas compatible avec les deux molécules par maille et un agrégat fer-soufre par protéine IRP1. En revanche, quatre sites sont prédictibles pour une forme triclinique et après plusieurs réindexations nous nous sommes aperçus que les cristaux 40 et 16 pouvaient être indexés en P1 et en $C222_1$ en plus du $P2_1$ et du C2 (Tableau III.5). A chaque nouveau traitement les statistiques obtenues étaient de bonne qualité (Tableau III.6).

Cristal	Grp	a (Å)	b (Å)	c (Å)	$\alpha(^{\circ})$	$\beta(^{\circ})$	$\gamma(^{\circ})$
16	$P2_1$	63,9	215,5	64,0	90	72	90
16	P1	64,0	64,1	215,6	90	90	108
16	$C222_1$	75,0	103,2	214,9	90	90	90
40	C2	103,5	75,1	214,7	90	90	90
40	P1	63,9	63,9	214,7	90	108	90
40	$C222_1$	75,1	103,5	214,7	90	90	90

Tableau III.5 : Paramètres de maille des différentes indexations possibles pour les cristaux 40 et 16.

Cristal - Grp d'espace	16 - C222 ₁	16 - P1	40 - C222 ₁	40 - P1
Ligne de lumière ESRF	ID14-4	ID14-4	ID14-4	ID14-4
Résolution	30-1,85	35-1,85	16-2,0	25-2,0
Nombre d'observations	545437	537827	355038	353799
Réflexions uniques	136857	471653	101337	187619
Complétude (%)	99,5 (98,0)	84,2 (60,4)	92,9 (51,4)	88,8 (52,0)
Redondance	4,0 (3,4)	1,1 (1,1)	3,5 (1,5)	1,9 (1,5)
$\langle I/I \rangle$	8,6 (3,9)	6,2 (2,6)	7,7 (3,6)	10,3 (4,5)
R _{sym} (%)	10,8 (31,9)	6,6 (22,6)	12,0 (18,6)	4,8 (13,4)
R _{meas} (%)	12,4 (37,7)	9,3 (32,0)	14,0 (25,0)	6,7 (19,0)

Tableau III.6 : Statistiques des jeux natifs après indexation dans les nouveaux groupes d'espace.

A ce stade de l'avancée des travaux nous n'étions pas en mesure de trancher de façon définitive sur le groupe d'espace applicable aux différents cristaux. Prenant en compte les groupes de plus haute symétrie nous avons essayé de localiser la position des super atomes de fer dans la structure mais sans grand succès. En effet, la protéine IRP1 a une masse moléculaire de 98 kDa pour 889 résidus ce qui est considérable par rapport au pouvoir de phasage des seuls agrégats. Nous nous sommes alors lancés dans la recherche de dérivés d'atomes lourds mais, pour cela, nous avons dû repartir sur des tests de cristallogenèse car les cristaux obtenus pour ces premières expériences n'ont jamais pu être reproduits.

2 Résolution du problème de la phase

2.1 Cristallisation

Fort de ces premiers résultats, certes frustrants mais très encourageants, nous sommes repartis de la condition 41 du *Crystal Screen I* (Hampton research) afin d'effectuer, à partir d'un nouveau lot d'échantillon de protéine, la recherche des nouvelles conditions de cristallisation. La protéine, toujours à la concentration de 20 mg/ml dans un tampon contenant 10 mM Hepes pH 7,4, a donné des cristaux dans une solution de 100 mM Tris-HCl pH 8,1, 5% isopropanol, 17% PEG 4000 et 2 mM Na-dithionite. Le problème pour nous a été de voir que ces cristaux évoluaient au cours du temps vers une lente dégradation pour finir, sous cinq jours, par être complètement inexploitable (Figure III.5). Nous avons donc cherché dans un

premier temps à stabiliser ces cristaux avant d'envisager l'élimination de leur croissance en agrégat.



Figure III.5 : Evolution au cours du temps de la morphologie des cristaux. Il y a une dégradation des cristaux sous cinq jours (j dans la figure).

Nous avons testé plusieurs solvants organiques afin de palier à ce problème mais aussi afin de nous débarrasser de l'isopropanol. En effet, comme toutes les expériences de cristallogénèse sont effectuées en boîte à gants, une étape de dégazage des solutions est nécessaire avant leur utilisation pour faire les gouttes. L'isopropanol est un solvant organique volatil et donc le simple fait de faire buller l'atmosphère de la boîte à gants dans la solution mère le contenant, afin d'éliminer toute trace d'oxygène, ne nous permet pas de contrôler sa concentration finale exacte. Ceci représente donc un désavantage majeur pour obtenir de nouveaux cristaux. Nous avons donc exploré différentes conditions tournant autour de la précédente en utilisant à chaque fois 5% de différents solvants. La condition, qui nous a permis de comparer les résultats obtenus, était constituée de base par 100 mM Tris-HCl pH 8,1, 16% PEG 4000 et 2 mM Na-dithionite auxquels était ajouté 5% de différents solvants (Figure III.6).

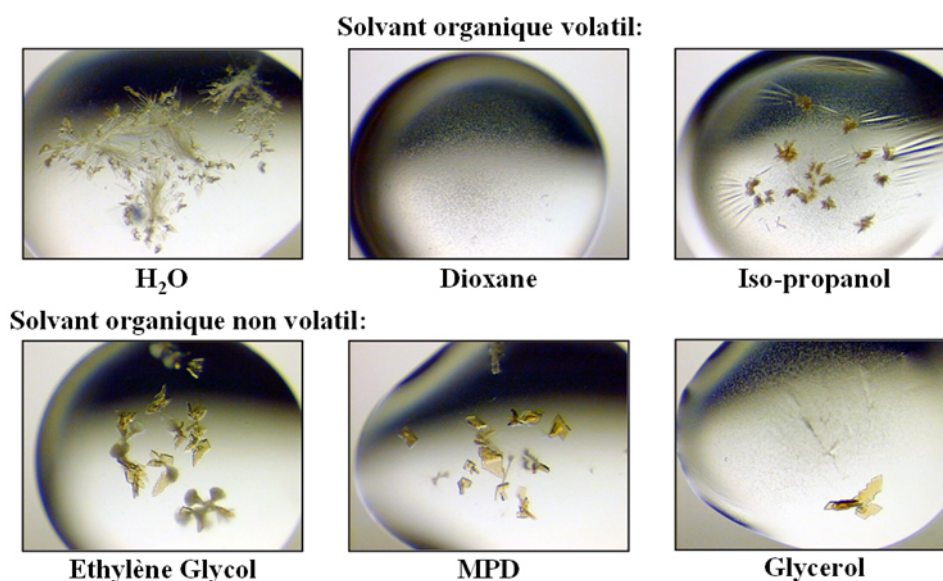


Figure III.6 : Stabilisation des cristaux avec différents solvants.

Après avoir choisi le glycérol comme agent stabilisateur, nous avons orienté nos efforts sur le faciès des cristaux. Dans cette optique, la concentration de la solution de protéine a été revue à la baisse et nous avons choisi de travailler à 10 mg/ml au lieu des 20 utilisés jusqu'à présent. Afin de contrôler le phénomène de nucléation nous avons aussi opté pour la technique d'ensemencement à partir des cristaux déjà obtenus. Une dilution sériée a permis de contrôler le nombre de noyaux de nucléation et donc le nombre de cristaux mais sans résoudre le phénomène d'agrégation (Figure III.7).

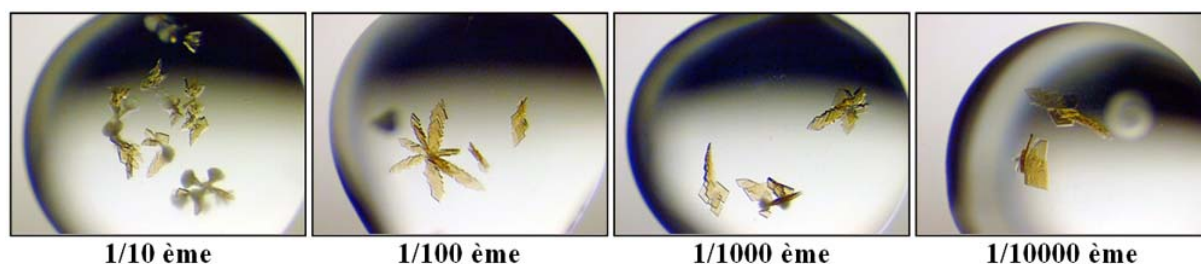
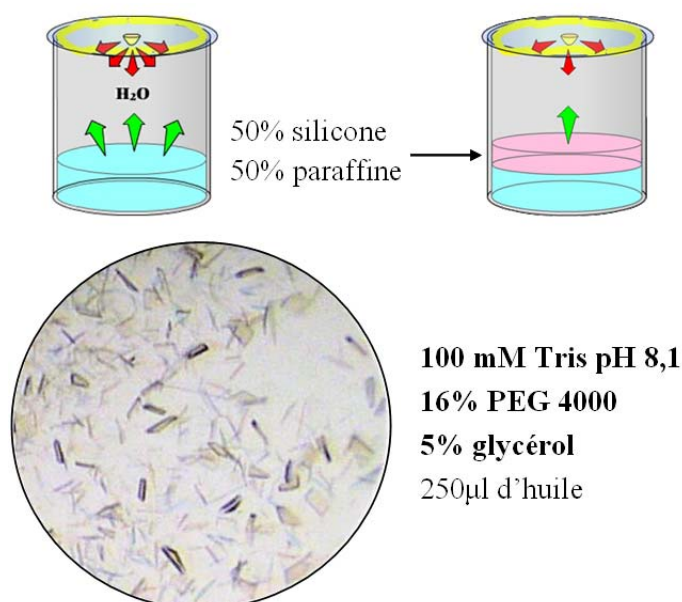


Figure III.7 : Rôle de la dilution sériée dans l'ensemencement sur le contrôle du nombre de noyau de nucléation.

Le problème a été résolu en utilisant une technique décrite pour la première fois par N Chayen (1997) qui a eu l'idée de limiter la diffusion de vapeur d'eau en plaçant, entre le réservoir et la goutte, une barrière composée d'un mélange d'huile minérale, comme la paraffine, et d'huile de synthèse, comme le silicone. Cette technique combinée à l'ensemencement nous a permis d'obtenir finalement des monocristaux d'une taille approximativement de 50 μm sur 30 μm sur 15 μm (Figure III.8).

Figure III.8 : Limitation de la diffusion de vapeur. L'emploi d'une barrière de silicone et de paraffine a permis de laisser le temps aux cristaux de pousser de façon plus homogène. Différents cocktails d'huiles ont été utilisés ainsi que différents volumes à déposer à la surface du réservoir.



2.2 Dérivés d'atomes lourds

La recherche des dérivés d'atomes lourds a été réalisée uniquement par trempage des cristaux natifs dans des solutions contenant le sel d'atome lourd en plus des agents

précipitants. Ces derniers ont été ajoutés en concentration suffisante pour prévenir la dissolution des cristaux pendant le temps du trempage. Les sels d'atomes lourds utilisés sont présentés dans le tableau III.7. Les temps de trempage ont été de 24 h dans tous les cas. Les cristaux ont ensuite été congelés dans la boîte à gants (voir § II.1.5.1). Tous ces dérivés ont été testés sur la ligne de lumière FIP-BM30a à l'ESRF. Après avoir analysé les premiers résultats seul le sel d'or s'était fixé à la protéine.

Nature du sel d'atome lourd	Concentration utilisée
K_2PtCl_4	5 et 10 mM
$KAu(CN)_2$	10 mM
$UO_2(CH_3COO)_2$	5 et 10 mM
$HgCl_2$	100 μ M
ErCl	10 mM
Gd DOTMA	100 mM
Gd DTPA-BMA	50 mM
NaBr	1 M
$IrCl_3$	10 mM

Tableau III.7 : Sels d'atomes lourds testés et concentrations utilisées.

2.3 Phasage par la méthode MAD

2.3.1 Mesure de fluorescence et choix des longueurs d'onde

L'utilisation du détecteur de fluorescence résolu en temps de la ligne FIP nous a permis de valider la présence d'or dans notre cristal d'IRP1. La longueur d'onde à laquelle nous avons décidé de travailler pour avoir un signal anomal à l'or était de 1,040 Å. A cette énergie nous nous attendions à observer, dans notre spectre de fluorescence, outre le signal de l'or celui des agrégats Fe-S et nous avons eu la surprise de trouver un pic supplémentaire non prédit qui correspondait au zinc (Figure III.9). A ce stade des recherches nous avons non plus deux diffuseurs anomaux mais trois. Des expériences supplémentaires nous ont permis de confirmer la présence de zinc au sein des cristaux natifs. Nous avons donc décidé de collecter un jeu de données SAD complet au seuil de l'or (cristal 1) et sur un cristal natif d'enregistrer un jeu de données MAD au seuil du fer ainsi qu'un jeu SAD au pic du zinc (cristal 2).

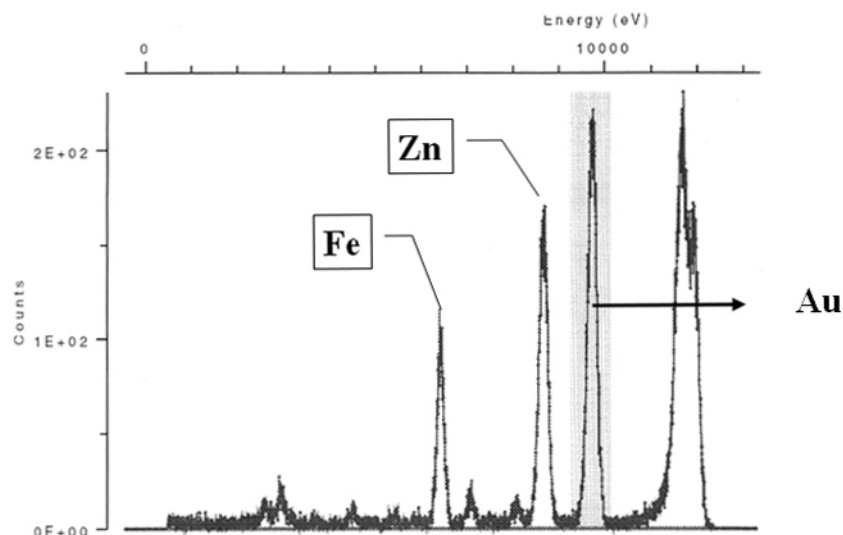


Figure III.9 : Spectre de fluorescence au seuil d'excitation de l'or. Le pic à droite correspond au signal diffusé.

Les différentes longueurs d'onde nécessaires à la réalisation de ces enregistrements ont toutes été choisies après l'acquisition des spectres de fluorescence au seuil d'excitation des diffuseurs anomaux, c'est-à-dire au seuil K du fer $\lambda=1,743 \text{ \AA}$, au seuil LIII de l'or $\lambda=1,039 \text{ \AA}$ et enfin au seuil K du zinc $\lambda=1,283 \text{ \AA}$ (Figure III.10).

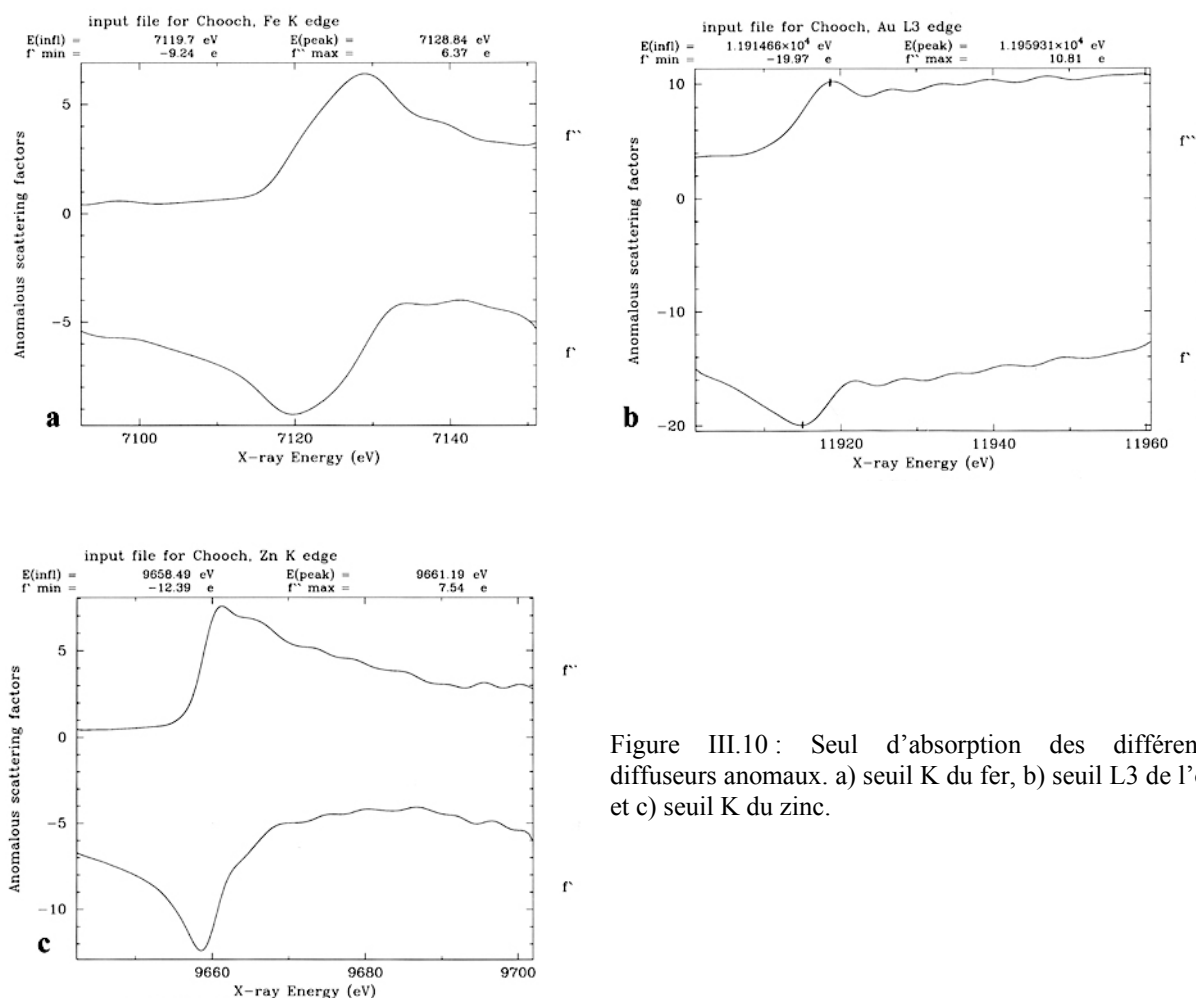


Figure III.10 : Seuil d'absorption des différents diffuseurs anomaux. a) seuil K du fer, b) seuil L3 de l'or et c) seuil K du zinc.

Les longueurs d'ondes, les énergies et les facteurs f'/f'' sont résumés dans le tableau III.8.

Collectes	Energie (eV)	λ (Å)	f'	f''
Fe maximum de f''	7128,8	1,739	-6,19	6,38
Fe minimum de f'	7119,7	1,741	-9,25	2,87
Au maximum de f'	11919	1,040	-17,48	10,20
Zn maximum de f''	9661,1	1,283	-9,69	7,50

Tableau III.8 : Paramètres des différentes longueurs d'onde utilisées lors des collectes et de leurs facteurs f' et f'' mesurés.

2.3.2 Intégration et mise à l'échelle – statistiques

L'indexation des données a été réalisée par le programme XDS (Kabsch, 1993) et les paramètres de maille sont résumés dans le tableau III.9.

Cristal	Grp	a (Å)	b(Å)	c(Å)	$\alpha(^{\circ})$	$\beta(^{\circ})$	$\gamma(^{\circ})$
1 - Au	C222 ₁	75,8	103,2	225,9	90	90	90
2 - max f''	C222 ₁	75,2	103,1	225,8	90	90	90
2 - min f''	C222 ₁	75,1	102,9	225,4	90	90	90
2 - Zn	C222 ₁	75,3	103,3	225,9	90	90	90

Tableau III.9 : Paramètres de maille des deux cristaux 1 et 2 juste après l'indexation.

La première constatation a été que la maille, qui auparavant était du triclinique ou du monoclinique selon l'indexation, avait changé pour de l'orthorhombique et qu'aucune ambiguïté n'était possible sur le groupe d'espace C222₁. La comparaison avec les jeux obtenus à partir de X16 et X40 n'était guère possible car, même si l'indexation en C222₁ était prise en considération, le paramètre c de la maille était différent de la nouvelle maille de 10 Å.

Cristal	1 - Au	2 - Fe max f''	2- Fe min f'	2- Zn
Ligne de lumière ESRF	FIP-BM30a	FIP-BM30a	FIP-BM30a	FIP-BM30a
Résolution	47-2,5	25-2,89	25-2,88	47-2,54
Nombre d'observations	448864	128429	115555	187987
Réflexions uniques	59196	36128	35331	53181
Complétude (%)	99,9 (99,8)	94,5 (71,7)	92,9 (68)	94,6 (72,9)
Redondance	7,6 (7,4)	3,6 (1,4)	3,3 (1,2)	3,5 (1,3)
$\langle I/I \rangle$	13,3 (5,9)	10,7 (2,7)	7,7 (1,6)	11,1 (3,3)
R_{sym} (%)	12,5 (34,8)	9,5 (33,6)	13,6 (55,7)	7,9 (23,1)
R_{meas} (%)	13,5 (37,5)	11,1 (44,4)	16,1 (74,4)	9,2 (30,6)

Tableau III.10 : Statistiques des jeux MAD et SAD collectés sur les cristaux 1 et 2.

Il était à noter aussi que les cristaux 1 et 2 ne présentaient qu'un faible degré de non isomorphisme, preuve que dans notre cas, l'utilisation d'un dérivé or n'avait pas beaucoup perturbé l'empilement cristallin. Le tableau III.10 présente les principales statistiques après traitement.

2.3.3 Nombre de molécules par unité asymétrique

De la même manière que pour les cristaux 40 et 16 (voir § III.1.2.2), nous avons calculé le coefficient de Matthews (Matthews, 1968) pour ces nouveaux cristaux. Le groupe d'espace $C222_1$ possède huit positions équivalentes et le volume de maille, quelque fut le cristal, était proche de 1767116 \AA^3 . Le V_m ainsi obtenu avait une valeur de $2,2 \text{ \AA}^3/\text{Da}$ pour une seule molécule d'IRP1 par unité asymétrique.

2.3.4 Détermination de la position des agrégats et des atomes lourds

L'ensemble des données ont été traitées avec la suite de programme de CCP4 (CCP4, 1994) et les cartes de Patterson nous ont permis de confirmer ce à quoi nous nous attendions à savoir, un site pour l'agrégat $[4\text{Fe-4S}]$, un site pour le zinc et un site pour l'or (Figure III.11 a, b et c).

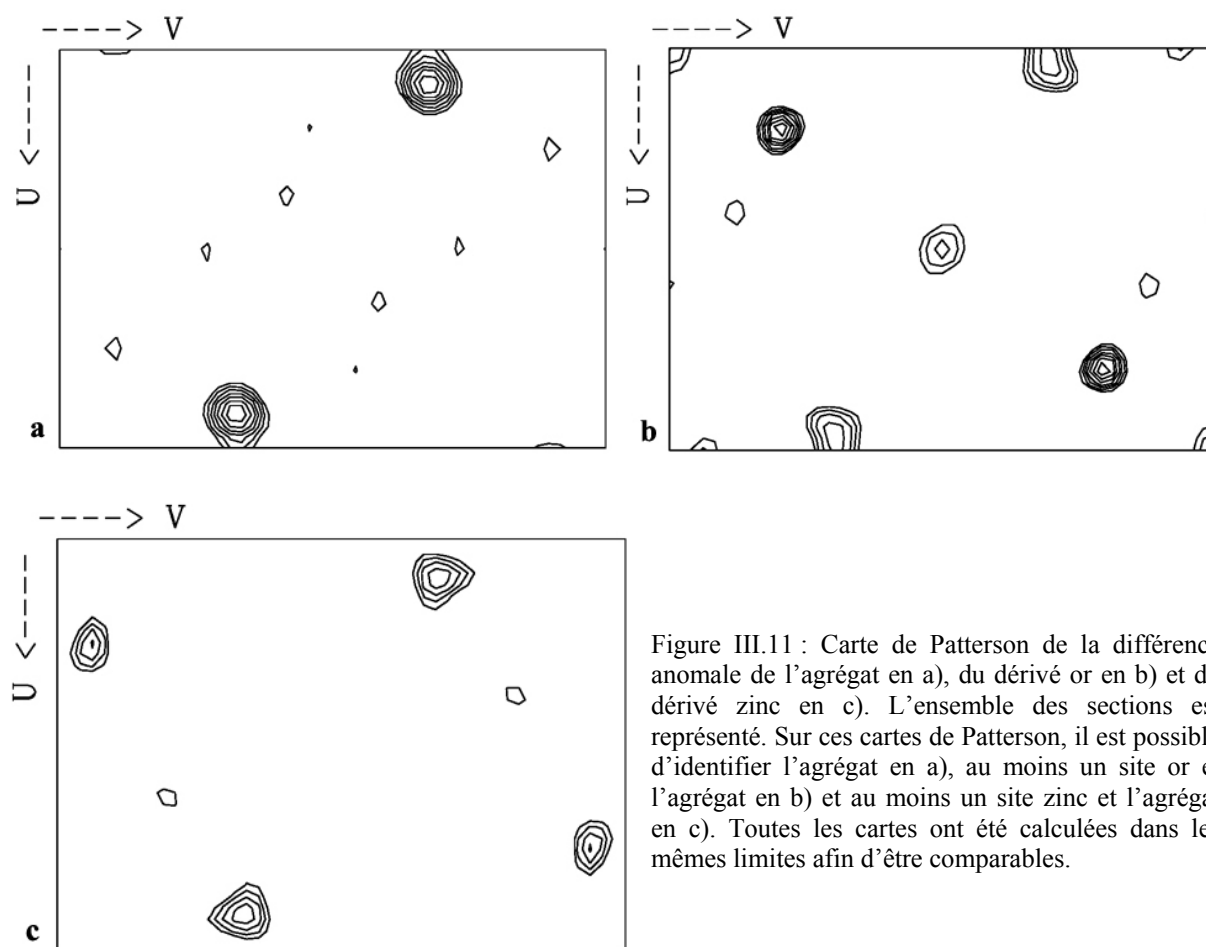


Figure III.11 : Carte de Patterson de la différence anormale de l'agrégat en a), du dérivé or en b) et du dérivé zinc en c). L'ensemble des sections est représenté. Sur ces cartes de Patterson, il est possible d'identifier l'agrégat en a), au moins un site or et l'agrégat en b) et au moins un site zinc et l'agrégat en c). Toutes les cartes ont été calculées dans les mêmes limites afin d'être comparables.

Là encore nous avons considéré l'agrégat comme un super atome de fer. Nous avons donc trois sites d'atomes lourds à déterminer et cette recherche a été réalisée à l'aide du programme SHELXD (Schneider et Sheldrick, 2002) qui, par des méthodes directes basées sur des Pattersons dans l'espace réciproque et la déconvolution des pics dans l'espace réel, est capable de localiser un grand nombre de diffuseurs anomaux à partir de données SAD ou de données MAD. En effet, bien que les méthodes directes ont été développées pour les petites molécules qui requièrent des résolutions atomiques inférieures à 1,2 Å, elles demeurent encore efficaces pour trouver les positions des atomes lourds car la distance les séparant est en générale plus grande que la résolution des données. Le programme a été utilisé sur le jeu obtenu pour le dérivé or car la redondance est le critère le plus important, bien devant la complétude et de la résolution (Schneider et Sheldrick, 2002). La partie qui nous intéressait, est présentée dans la figure III.12.

PATFOM 32.70

x	y	z	sof	height
0.79941	0.19913	0.03103	1.000	99.90
0.22518	0.07639	0.11040	1.000	43.57
0.56527	0.01296	0.23960	1.000	36.99
0.29427	0.15649	0.22256	1.000	32.17
0.20346	0.29677	0.03001	1.000	11.72
0.69685	0.66098	0.02956	1.000	10.45
0.78001	0.25447	0.16131	1.000	5.96

Minimum distances (top row, 0 if special position) and PSMF (bottom row)

Peak	x	y	z	self	cross-vectors
99.9	0.7994	0.1991	0.0310	41.8 134.1	A1
43.6	0.2252	0.0764	0.1104	52.3 28.1	39.0 43.9 A2
37.0	0.5653	0.0130	0.2396	11.0 17.0	53.9 29.1 38.0 7.8 A3
32.2	0.2943	0.1565	0.2226	33.6 5.0	56.3 27.2 20.1 46.1 13.4 4.5
11.7	0.2035	0.2968	0.0300	41.4 123.6	15.6 29.2 53.4 46.4 35.0 2.9 0.0 0.0

10.5	0.6968	0.6610	0.0296	35.8 0.0	21.4 20.4 57.1 44.2 14.0 17.7 0.0 0.0 0.0 0.0
6.0	0.7800	0.2545	0.1613	52.1 0.0	30.0 35.4 34.6 28.7 43.9 44.4 7.2 0.0 0.0 0.0 2.8 0.0

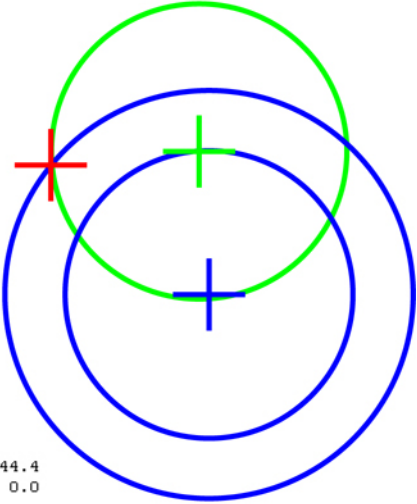


Figure III.12 : Fichier de sortie de SHELXD. Les cercles ont des rayons correspondant aux distances séparant les positions des atomes lourds. Voir texte pour les explications.

Un score est attribué aux résultats sous la dénomination de PATFOM pour *Patterson Figure Of Merit*. Ce score permet de faire un tri rapide et ensuite une analyse plus détaillée est nécessaire. Les coordonnées des sites proposés sont classées puis pour chaque paire d'atomes deux valeurs lui sont attribuées. Le nombre du haut correspond à la distance

les séparant, par exemple 39 Å entre A2 et A1 et 53,9 Å entre A3 et A1, en tenant compte des symétries du groupe d'espace. Le nombre du bas indique le PMF, *Patterson Minimum Function* (Nordman, 1966), calculé pour tous les vecteurs entre deux atomes ce qui correspond aux vecteurs croisés ou *cross-vectors*. La première colonne est basée sur les vecteurs propres ou *self-vectors*, c'est-à-dire les vecteurs entre l'atome et ses symétriques. Nous n'avons considéré, dans un premier temps, que les trois premières solutions en supposant que le premier site devait être celui de l'or, le deuxième celui de l'agrégat fer soufre et le dernier celui du zinc.

2.3.5 Phasage et carte de densité électronique

Nous avons commencé à affiner ce modèle à l'aide du programme SHARP (de la Fortelle et Bricogne, 1997). L'affinement a été réalisé de façon progressive en dissociant tous les jeux obtenus pour les différents dérivés avant de tous les combiner pour le phasage final. SHARP (de la Fortelle et Bricogne, 1997) est actuellement le seul programme de phasage à intégrer complètement le maximum de vraisemblance à ses procédures de calcul. Parmi les outils que propose SHARP (de la Fortelle et Bricogne, 1997), la carte résiduelle permet de vérifier les erreurs de positionnement des atomes lourds utilisés lors du phasage mais aussi, et surtout, les sites minoritaires présents dans la structure. C'est ainsi que lors du traitement du jeu SAD du dérivé or, nous avons pu identifier un site supplémentaire. Les coordonnées de ce nouveau site se sont avérées avoir été prédites par SHELXD (Schneider et Sheldrick, 2002) comme étant le quatrième site de la liste. L'ensemble des jeux a été utilisé pour résoudre le problème de la phase. Un facteur B de 80 Å² a été utilisé pour l'agrégat fer soufre afin de mimer l'étalement de la densité électronique dû à sa taille. Un résumé des statistiques de phasage est présenté dans le tableau III.11.

Jeux	Réflexions centrées			Totalité des réflexions		
	R _{Cullis}	P.P.	FOM	R _{Cullis}	P.P.	FOM
1 - Au	0,675 (0,877)	1,147 (0,762)		0,597 (0,869)	1,590 (0,948)	
2 - Fe maximum de f''	- (-)	- (-)		- (-)	- (-)	
2 - Fe minimum de f'	0,862 (0,527)	0,286 (0,085)	0,510 (0,363)	0,826 (0,555)	0,318 (0,139)	0,586 (0,451)
2 - Zn	0,675 (0,615)	0,530 (0,137)		0,602 (0,567)	0,635 (0,223)	

Tableau III.11 : Statistiques du phasage en sortie de SHARP (de la Fortelle et Bricogne, 1997) pour un modèle complet avec quatre diffuseurs anomaux. Les valeurs entre parenthèses correspondent à la dernière coquille de résolution. La longueur d'onde collectée sur le cristal 2 au maximum de f'' a servi de référence ce qui explique l'absence de valeurs. P.P. pouvoir de phasage, FOM figure de mérite.

Les phases ainsi obtenues n'étaient pas d'une très bonne qualité mais les premières cartes de densité électronique permettaient de distinguer les zones de protéine des zones de solvant (Figure III.13a). Cette première carte restait tout de même non interprétable pour la construction du modèle. La connaissance du contenu de la maille nous a permis de faire une estimation du pourcentage de solvant du cristal aux environs de 44%. Nous avons alors réalisé une étape d'aplatissement de solvant à l'aide du programme SOLOMON (Abrahams et Leslie, 1996). Cette étape de modification de densité a amélioré la qualité des cartes de densité électronique, notamment en augmentant le contraste entre les zones de protéine et les zones de solvant (Figure III.13b).

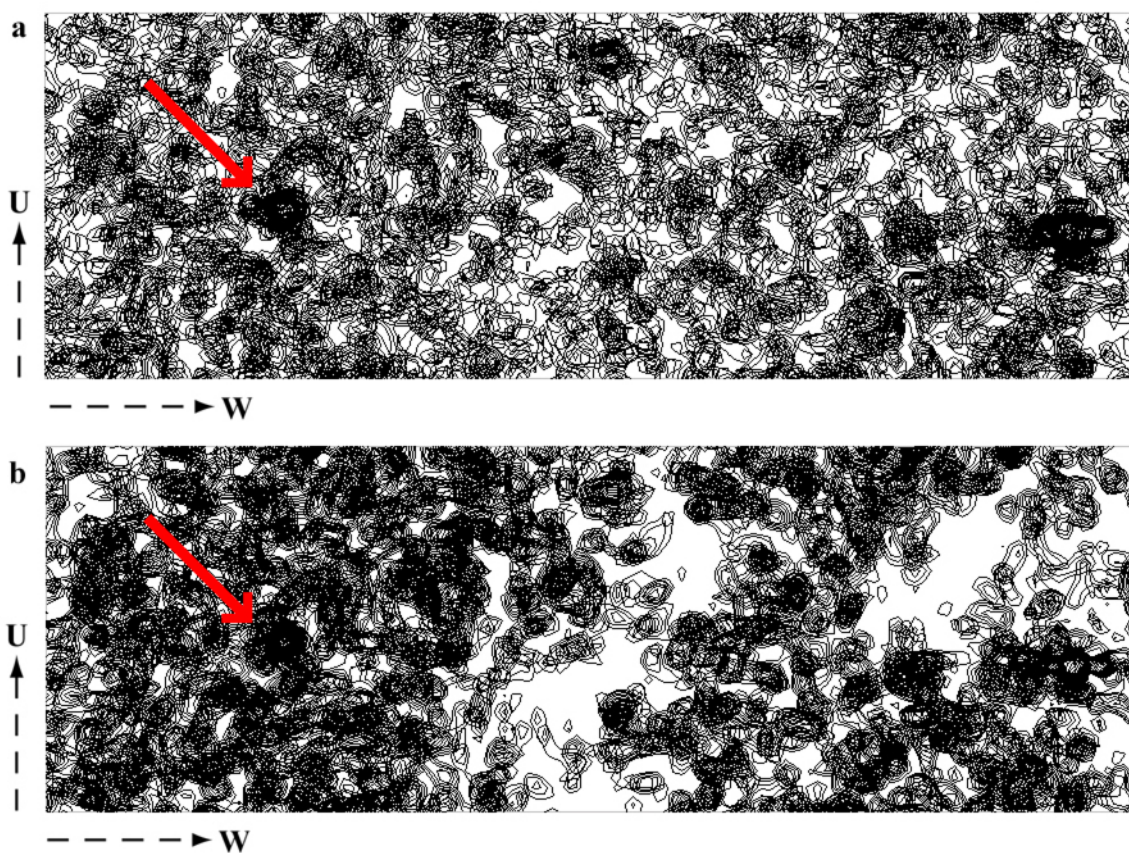
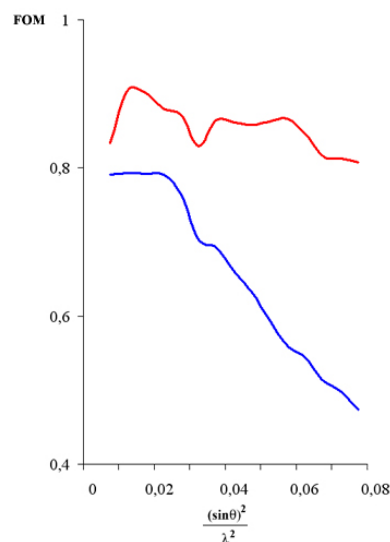


Figure III.13 : Projection suivant l'axe b de la densité électronique. a) Densité électronique obtenue avec les phases issues de SHARP (de la Fortelle et Bricogne, 1997) et b) densité électronique obtenue après aplatissement de solvant avec SOLOMON (Abrahams et Leslie, 1996). La position de l'agrégat dans ces cartes de densité électronique est indiquée par les flèches rouges.

La figure III.14 montre l'évolution de la figure de mérite (pondération accordée à chaque facteur de structure en fonction de la qualité de la phase) en fonction de la résolution avant et après aplatissement de solvant. La carte ainsi calculée s'est révélée être de bonne qualité et interprétable malgré certaines zones mal définies.

Figure III.14 : Evolution de la figure de mérite en fonction de la résolution. La courbe en bleu correspond à l'évolution de la figure de mérite associée aux phases déterminées par SHARP (de la Fortelle et Bricogne, 1997) et celle en rouge, à la sortie de SOLOMON (Abrahams et Leslie, 1996).



```
Total residues placed: 32 of 840 or 3%
Residues built without side chains: 608
Total residues built: 640 or 76%
```

```
Total score for this arrangement: 2.7
```

```
Build score = residues built + placed - 2* # of chains not
assigned to sequence = 486.
```

```
Match of sequence to side-chain density:
```

RESIDUE	CC	Z-SCORE	--SIDE CHAIN DENSITY/RMS--		
			AVG	MIN	MAX
107 LEU	0.57	1.73	1.44	0.20	2.65
108 GLY	0.45	1.37			
109 GLY	0.44	1.35			
110 ASP	0.43	0.81	0.99	0.67	1.38
111 PRO	0.36	1.05	0.99	-0.11	1.97
112 GLU	0.35	0.17	0.92	-0.82	2.50
113 LYS	0.39	0.59	1.26	0.45	1.71
114 ILE	0.27	0.12	0.80	0.48	1.11
115 ASN	0.39	0.56	0.96	-0.02	1.95
116 PRO	0.41	1.37	1.73	1.16	2.18

RESIDUE	CC	Z-SCORE	--SIDE CHAIN DENSITY/RMS--		
			AVG	MIN	MAX
279 GLU	0.43	0.83	1.56	0.00	3.00
280 PHE	0.27	0.14	0.59	-0.10	1.56
281 PHE	0.26	0.01	0.27	-0.44	1.19
282 GLY	0.43	1.31			
283 PRO	0.39	1.25	0.86	0.65	1.23
284 GLY	0.35	0.91			
285 VAL	0.52	1.50	0.85	-0.12	2.42
286 ALA	0.33	0.40	0.55	0.55	0.55
287 GLN	0.47	1.31	1.08	-0.20	2.15
288 LEU	0.62	2.03	1.07	0.22	1.66
289 SER	0.60	1.48	3.00	2.29	3.72

RESIDUE	CC	Z-SCORE	--SIDE CHAIN DENSITY/RMS--		
			AVG	MIN	MAX
476 GLY	0.68	2.52			
477 VAL	0.32	0.50	0.33	-0.35	0.75
478 VAL	0.53	1.56	1.40	0.04	2.99
479 THR	0.37	0.75	0.72	0.36	1.08
480 TYR	0.28	0.19	-0.10	-0.91	0.55 TRUNCATED AT CB
481 TYR	0.31	0.46	0.66	-0.33	1.78
482 LEU	0.56	1.69	1.15	0.96	1.49
483 GLN	0.38	0.58	0.61	-0.37	2.87
484 GLU	0.46	1.05	1.37	-0.10	2.56
485 SER	0.58	1.40	1.98	1.98	1.98
486 GLY	0.42	1.23			

	CC	Z-SCORE	RHO-AVG	N
MEAN:	0.43	1.01	1.04	32

Figure III.15 : Extrait du fichier de sortie du programme RESOLVE (Terwilliger, 2000). Une liste des acides aminés présents dans la séquence de la protéine et placés dans la densité se trouve aussi associée à ce fichier.

Le programme RESOLVE (Terwilliger, 2000), qui utilise la théorie du maximum de vraisemblance, a été utilisé avec la carte de densité ainsi obtenue. A chaque cycle de construction, les nouvelles phases sont combinées aux phases expérimentales ce qui nous a permis au final d'avoir une carte d'une grande qualité. Le résultat a été assez encourageant car le programme a réussi à placer 608 résidus d'alanine et de glycine dans la carte auxquels se sont ajoutés 32 acides aminés comme le montre le fichier de sortie du programme dans la Figure III.15.

L'aconitase mitochondriale ne pouvant pas nous servir pour faire du remplacement moléculaire, il était néanmoins clair, qu'au vu de l'activité isomérase d'IRP1 et de son appartenance à la famille des aconitases, le cœur catalytique de la protéine devait être proche de celui de l'aconitase. De plus, nous savions que les cystéines, qui sont impliquées dans la coordination de l'agrégat Fe-S, se trouvent dans une des zones les plus conservées d'IRP1, ce qui nous a laissé à penser que le repliement structural de cette région devait être relativement proche d'une structure à l'autre. Fort de la qualité de la carte de ce premier modèle, obtenu par construction automatique, et de la structure déjà existante de l'aconitase mitochondriale de bœuf, nous avons pu continuer la construction de la protéine IRP1 manuellement.

2.3.6 Construction du premier modèle

La localisation de la densité électronique de l'agrégat [4Fe-4S] dans la carte a permis de prépositionner la structure de l'aconitase mitochondriale (Figure III.16b). Les éléments de structure secondaire autour de l'agrégat étant conservés nous avons pu bénéficier, en nous appuyant sur le modèle construit par RESOLVE (Terwilliger, 2000) (Figure III.16c), d'une solide base pour envisager l'organisation et la construction du cœur de la protéine.

La superposition du modèle incomplet et de l'aconitase nous a permis de constater que le programme RESOLVE (Terwilliger, 2000) avait construit des portions de protéine sur deux molécules IRP1 reliées par symétrie. Ceci corrigé nous avons une carte bien continue dans laquelle il était relativement aisé de suivre le tracé de la chaîne principale sauf dans certaines zones encore mal définies. A partir du modèle, nous avons continué la construction mais cette carte présentait néanmoins peu d'informations sur les chaînes latérales. La construction a été effectuée à l'écran graphique à l'aide du programme TURBO (Roussel et al., 1989). L'ajout de fragments d'hélice et le remplacement des glycines et des alanines ont été suivis de calculs de minimisation d'énergie afin d'améliorer la géométrie du modèle, ainsi que d'affinements locaux dans l'espace réel de façon à mieux placer les résidus en densité.

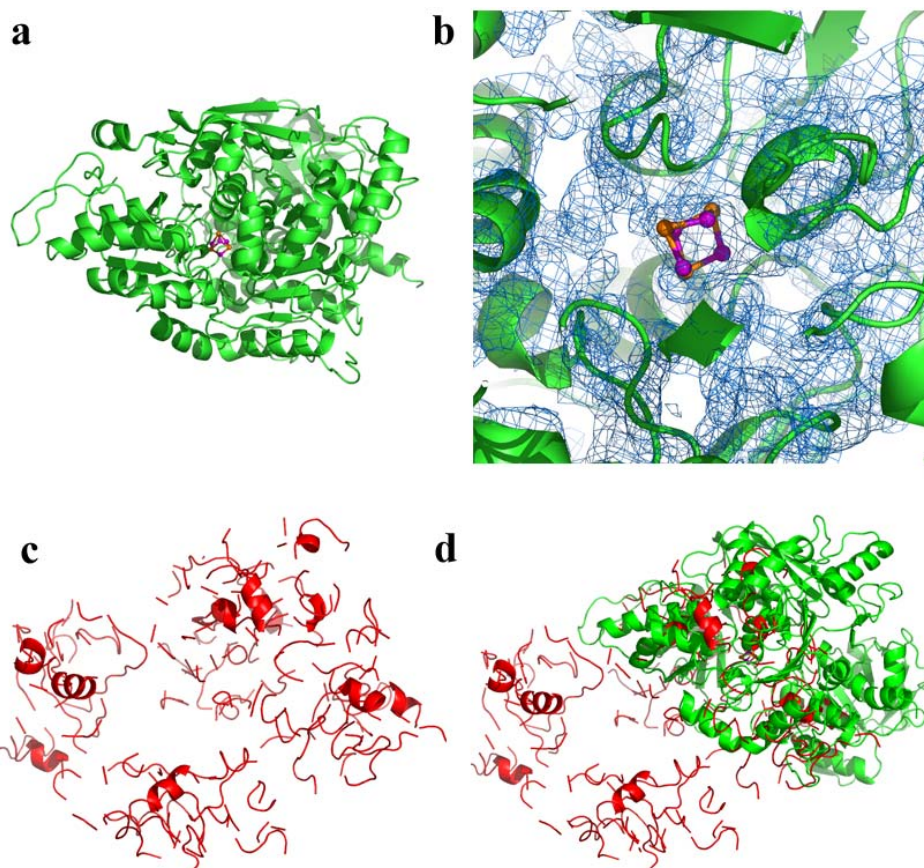


Figure III.16 : Obtention du premier modèle. a) Structure de l'aconitase mitochondriale de bœuf (code d'accès PDB : 1C96). b) Positionnement de l'agrégat [4Fe-4S] de l'aconitase et des structures secondaires de son environnement direct dans la carte de densité électronique. c) Modèle construit par le programme RESOLVE (Terwilliger, 2000). d) Superposition de la structure de l'aconitase et de notre modèle incomplet.

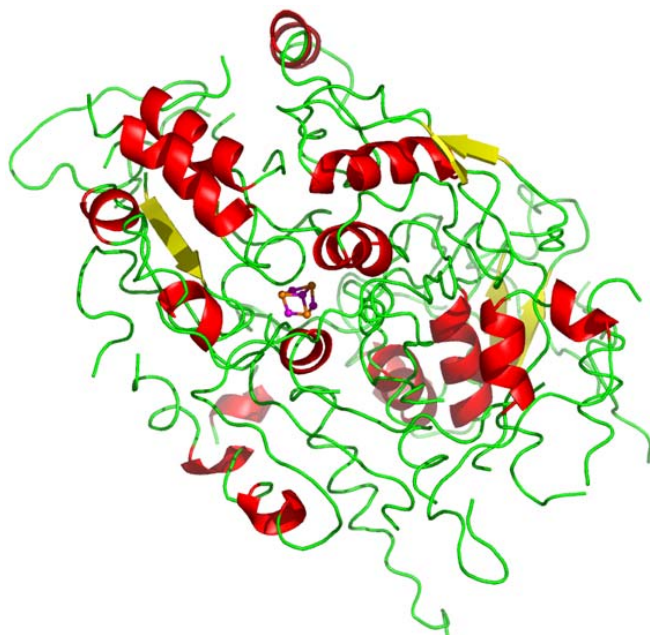


Figure III.17 : Structure incomplète d'IRP1.

Malgré nos efforts, nous avons rapidement été bloqués par le manque d'information dans certaines zones de la carte. Le modèle était complet à plus de 67%, 604 acides aminés placés sur les 889 que compte la protéine, mais les affinements successifs ne changeaient rien

à la situation. Le facteur R était de 45,9% et le R_{libre} de 47,4%, preuve que certaines parties de la protéine étaient encore mal construites (Figure III.17). Nous avons alors décidé de reprendre le phasage mais en utilisant une option propre au programme SHARP (de la Fortelle et Bricogne, 1997).

3 Résolution de la structure

3.1 Amélioration des cartes de densité électronique

SHARP (de la Fortelle et Bricogne, 1997) offre la possibilité d'utiliser les phases externes d'un modèle. Ce modèle peut être une solution de remplacement moléculaire ou alors un modèle incomplet construit à la main. L'avantage d'un tel système réside dans le fait que le programme va interpréter ces phases comme un système de pondération et non pas comme une base pour ses calculs évitant ainsi les problèmes de biais d'un modèle partiellement faux. Afin de pouvoir résoudre ce problème nous avons du extraire, du modèle précédent, les coefficient d'Hendrickson et Lattman (1970). Pour cela le programme SFTOOLS (CCP4, 1994) nous a été utile pour générer les coefficients A et B du modèle alors que les C et D sont égaux à 0 du fait de la répartition unimodale des phases. Nous avons réitéré le phasage sans rien changer aux paramètres utilisés ultérieurement. Seul la figure de mérite a radicalement évoluée et tout spécialement pour les données à « haute » résolution (Figure III.18).

a							b						
BINS	Dmin	Dmax	Nacen	FOMacen	Ncen	FOMcen	BINS	Dmin	Dmax	Nacen	FOMacen	Ncen	FOMcen
1	47.35	10.84	272	0.87140	140	0.61184	1	47.44	10.90	269	0.87843	139	0.65223
2	10.84	7.77	552	0.82723	136	0.66290	2	10.90	7.81	535	0.84085	136	0.67144
3	7.77	6.37	705	0.82691	120	0.62875	3	7.81	6.40	708	0.89032	117	0.77586
4	6.37	5.53	868	0.74657	119	0.54869	4	6.40	5.56	853	0.84521	117	0.72609
5	5.53	4.95	952	0.70942	117	0.41687	5	5.56	4.98	950	0.84832	117	0.77149
6	4.95	4.52	1071	0.63369	110	0.43068	6	4.98	4.55	1052	0.84925	114	0.76540
7	4.52	4.19	1182	0.57726	111	0.37873	7	4.55	4.21	1174	0.85842	103	0.85066
8	4.19	3.92	1269	0.55036	115	0.35311	8	4.21	3.94	1264	0.83074	123	0.83415
9	3.92	3.70	1357	0.50975	99	0.36680	9	3.94	3.72	1343	0.78867	99	0.80065
10	3.70	3.51	1458	0.47123	122	0.28334	10	3.72	3.53	1422	0.76210	118	0.71694
11	3.51	3.35	1519	0.44680	118	0.25298	11	3.53	3.37	1508	0.75861	114	0.72977
12	3.35	3.21	1615	0.42020	105	0.24179	12	3.37	3.22	1610	0.71610	98	0.64955
13	3.21	3.08	1728	0.66867	128	0.69988	13	3.22	3.10	1689	0.80583	125	0.83000
14	3.08	2.97	1740	0.75929	115	0.85463	14	3.10	2.99	1743	0.84120	118	0.90265
15	2.97	2.87	1834	0.71648	129	0.77267	15	2.99	2.89	1778	0.81280	125	0.85123
16	2.87	2.78	1866	0.69546	138	0.73317	16	2.89	2.79	1861	0.78154	137	0.77163
17	2.78	2.70	1971	0.67412	127	0.71461	17	2.79	2.71	1940	0.81631	122	0.80402
18	2.70	2.62	2009	0.60027	151	0.67991	18	2.71	2.63	2009	0.77470	148	0.79274
19	2.62	2.55	2049	0.45178	138	0.36328	19	2.63	2.56	2003	0.73282	130	0.69051
20	2.55	2.49	1812	0.20338	119	0.00000	20	2.56	2.50	2129	0.65049	150	0.58966
OVERALL			27829	0.58618	2457	0.51028	OVERALL			27840	0.78858	2450	0.75599

Figure III.18 : Récapitulatif des figures de mérite. a) FOM sans l'apport de phases externes et b) FOM en utilisant les coefficients d'Hendrickson et Lattman (1970) pour le phasage.

L'aplatissement de solvant a, par la suite, amélioré encore la qualité de la carte (Figure III.19).

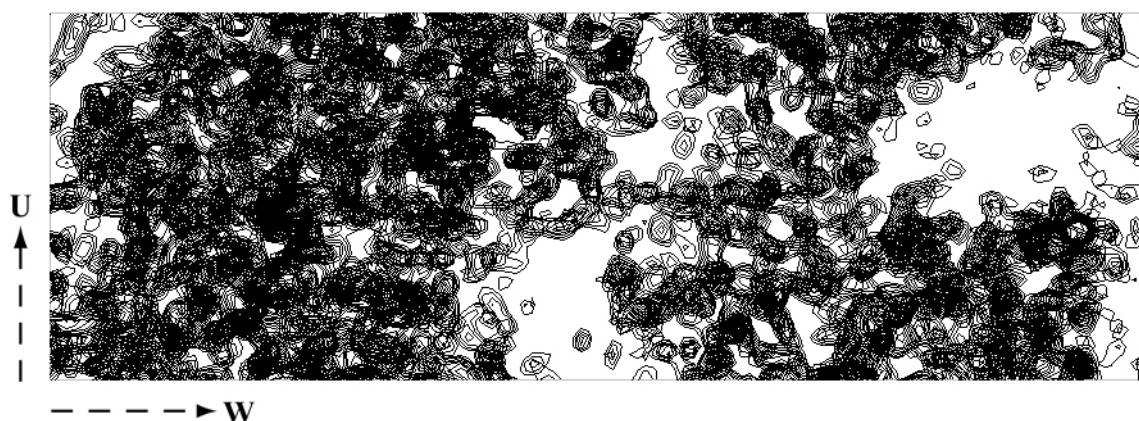


Figure III.19 : Projection suivant l'axe b de la densité électronique issue de SHARP (de la Fortelle et Bricogne, 1997) après l'utilisation de phases externes et aplatissement de solvant par SOLOMON (Abrahams et Leslie, 1996).

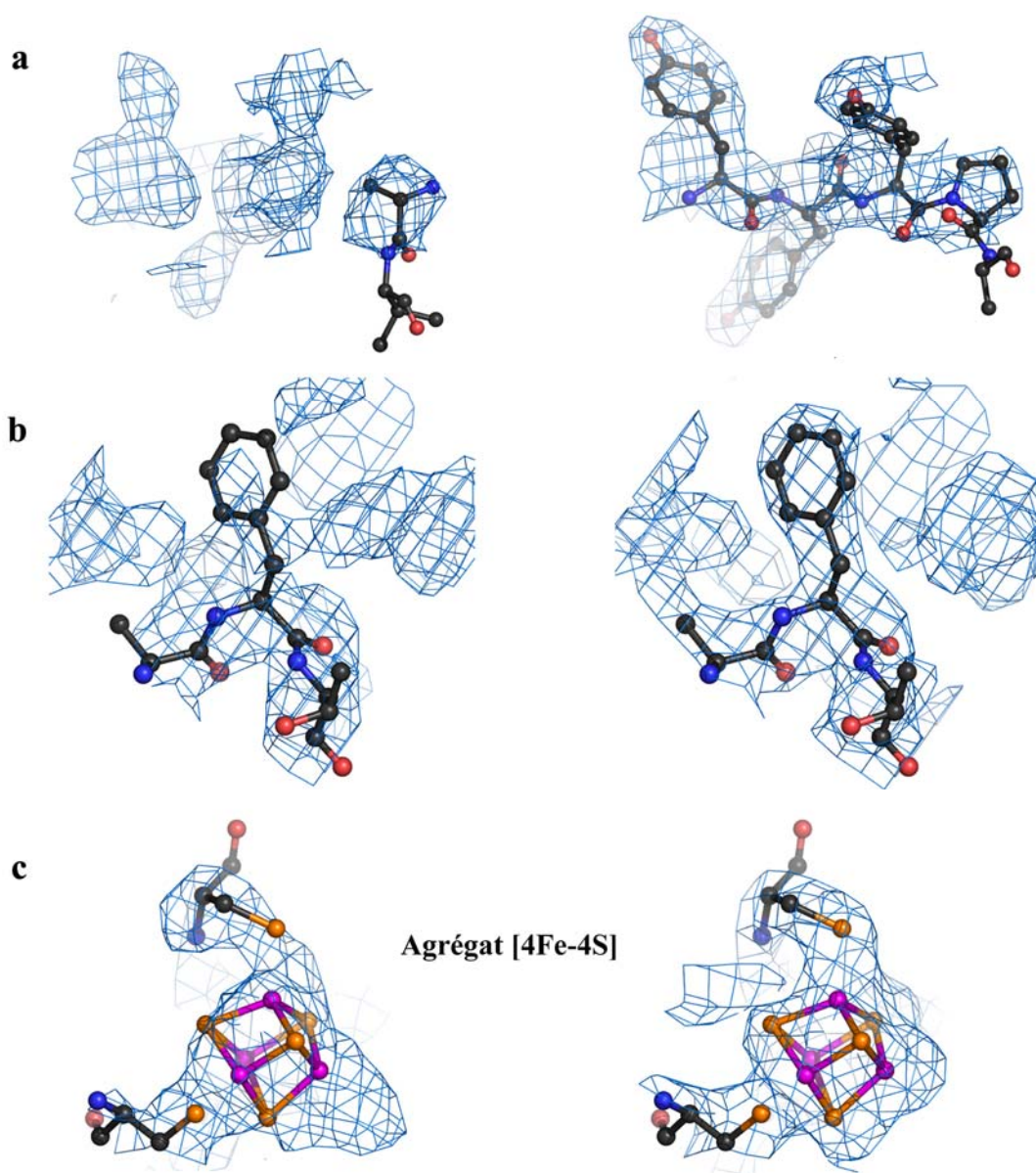


Figure III.20 : Amélioration de la carte de densité électronique. A gauche avant l'utilisation des phases externes et à droite après. a) Exemples avec la meilleure continuité de la chaîne principale, b) amélioration de la définition des chaînes latérales et b) meilleure définition de l'agrégat dans la densité.

Ainsi certaines régions de la densité électronique alors encore mal définies étaient maintenant interprétables. Le tracé de la chaîne principale était moins haché (Figure III.20 a) et la carte présentait plus d'informations sur les chaînes latérales que la précédente (Figure III.20 b). De même l'agrégat [4Fe-4S] était mieux défini dans la densité ce qui est illustré sur la Figure III.20 c. En nous appuyant sur ces nouveaux résultats nous avons décidé de reprendre la construction du modèle à l'écran graphique.

3.2 Nouveau modèle moléculaire

Le modèle a été construit à plus de 82% avec 736 acides aminés placés sur les 889 que contient la protéine (Figure III.21). Nous avons pu corriger, de la sorte, bon nombre d'erreurs dans notre modèle mais nous nous sommes rapidement heurtés au même problème que précédemment, à savoir, le manque d'information dans la densité électronique définissant certaines régions de la protéine plutôt exposées au solvant. Le facteur R était de 39% et le R_{libre} de 43%.

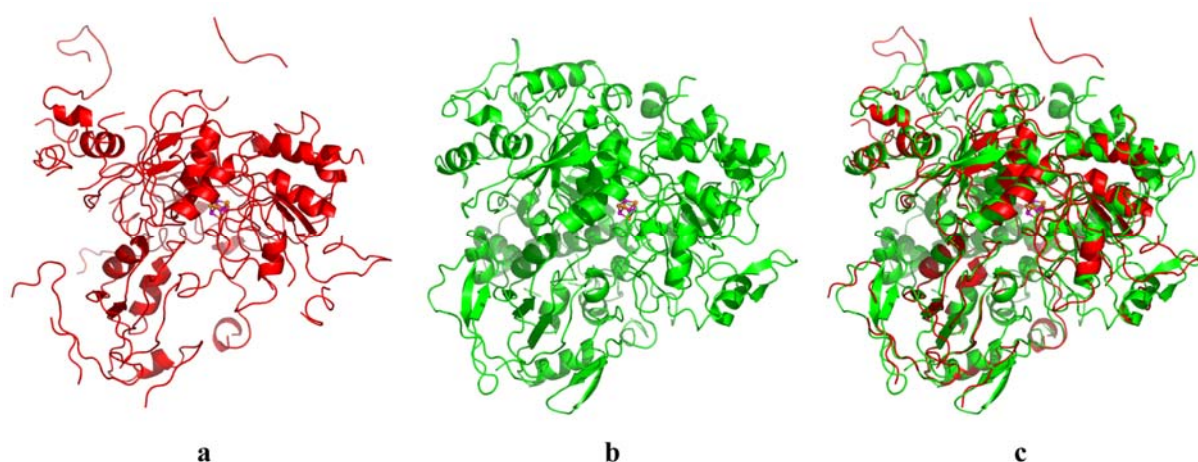


Figure III.21 : Amélioration du modèle. a) Premier modèle obtenu avec 604 AA placés. b) Nouveau modèle avec 736 AA placés. c) Superposition des deux modèles.

3.3 Remplacement moléculaire

3.3.1 Forme à haute résolution monoclinique $P2_1$

Afin de pouvoir nous absoudre des problèmes rencontrés lors de la construction avec la carte de densité électronique, nous avons décidé, à ce stade de nos investigations, d'essayer un remplacement moléculaire sur le jeu natif à plus haute résolution, à savoir le cristal 16. Nous avons aussi fait le choix d'utiliser les données du cristal 16 indexés dans le groupe d'espace de plus haute symétrie, $P2_1$.

Les essais de remplacement moléculaire avec comme modèle initial le premier modèle, trop incomplet et imprécis, n'avaient pas permis de dégager une quelconque solution cohérente, ni en rotation ni en translation. En revanche le second modèle nous a permis d'aboutir à une solution. Le remplacement moléculaire a été effectué avec le programme PHASER (Read, 2001). Afin d'accroître ses performances et d'augmenter les chances de trouver une solution, ce programme s'appuie sur la théorie du maximum de vraisemblance. Contrairement à AMORE (Navaza, 1994) et aux autres programmes de remplacement moléculaire, MOLREP (Vagin et Teplyakov, 1997)..., PHASER (Read, 2001) n'évalue pas les solutions qu'il trouve en terme de corrélation ni de facteur R mais il leur attribue une valeur dénommée *Z score*. Lorsque ce score est inférieur à 5, aucune solution n'a été trouvée ; entre 5 et 6, elle est improbable ; entre 6 et 7, elle est possible ; entre 7 et 8, elle est probable et au-delà de 8, elle est définitive. Dans notre cas une solution unique à deux corps, très contrastée en rotation et en translation, avec comme *Z score* respectivement 41 et 15, a été trouvée. Ce résultat confirmait bien notre hypothèse quant au contenu de la maille. Le programme nous a calculé directement les F_{calc} et ϕ_{calc} correspondant au modèle de la carte de densité électronique transposée dans la maille native du cristal 16. La nouvelle carte de densité électronique a été utilisée pour poursuivre la construction manuelle du modèle. Nous avons pu construire 826 des 889 acides aminés et ce troisième modèle avait un facteur R de 26% et un R_{libre} de 30% à 2,5 Å (Figure III.22).

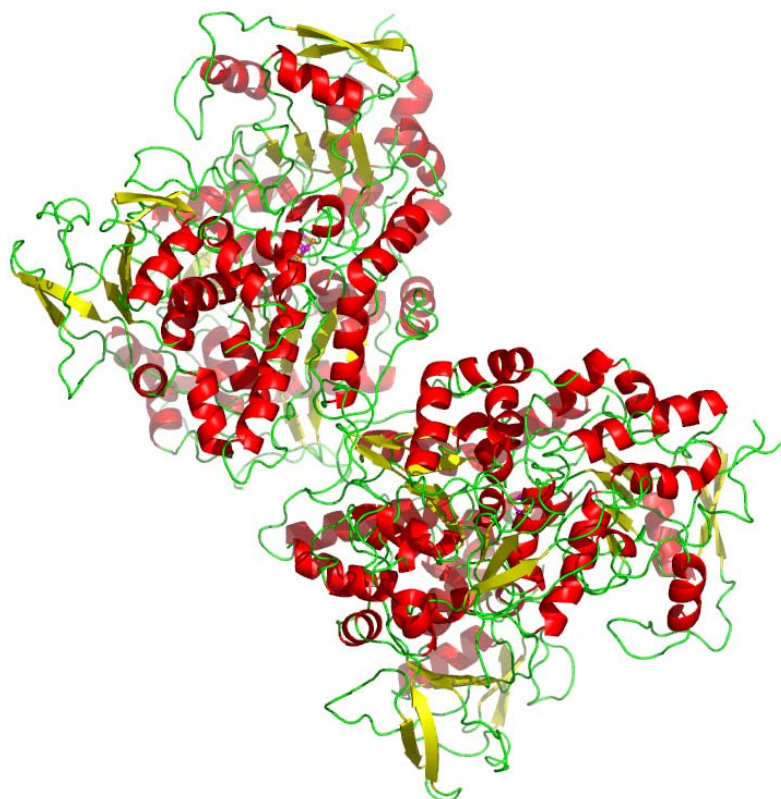


Figure III.22 : Structure incomplète d'IRP1 dans le cristal 16.

En effectuant cette opération nous avons gagné en résolution mais aussi en qualité. En effet, nous étions bloqués à 2,9 Å, qui était la résolution du jeu de données enregistré au maximum de f'' à la longueur d'onde du fer qui nous avait servi de référence pour le phasage, alors qu'avec le cristal 16 le modèle a été construit et affiné à 1,85 Å. Nous avons encore pu constater l'amélioration des cartes, au travers du fait que chaque atome de fer était maintenant défini dans la densité électronique, mais aussi par la définition des chaînes latérales dans la densité, surtout pour les acides aminés aromatiques (Figure III.23a et b). Malgré cela, il demeurait toujours des zones de densité hachée dans laquelle nous ne pouvions pas construire la chaîne polypeptidique (Figure III.23c).

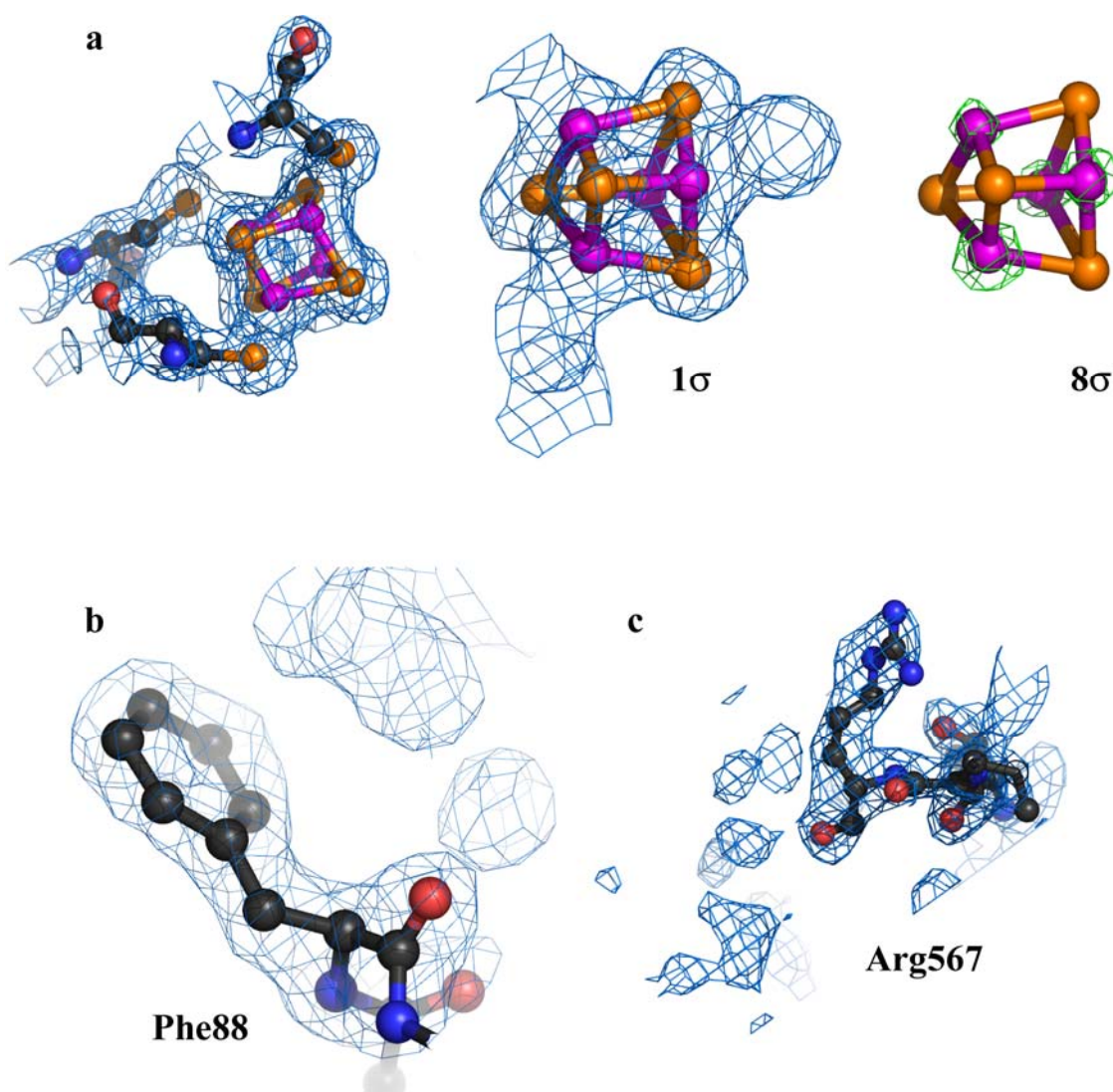


Figure III.23 : Suivi de l'évolution de la carte de densité électronique. a) L'agrégat est parfaitement relié au reste de la protéine du fait même de la continuité de la densité, de plus il est parfaitement défini dans la densité (seuil à 1σ) et chaque atome de fer est très bien localisé (seuil à 8σ). b) Les chaînes latérales des résidus aromatiques présentent des dépressions, voir des trous au centre du cycle, dans la densité électronique qui les définit. c) Région encore non interprétable de la densité.

3.3.2 Forme orthorhombique C222₁

Nous avons donc procédé avec une seule molécule du modèle P2₁ à un nouveau remplacement moléculaire sur le jeu de données du dérivé or qui avait une résolution de 2,5 Å. Cela n'a posé aucun problème, avec un *Z score* de 55 pour la rotation et de 29 pour la translation, et nous avons réussi à construire la totalité de la protéine dans la nouvelle carte de densité électronique. Même les régions non interprétables à haute résolution ont été résolues sans difficulté (Figure III.24)

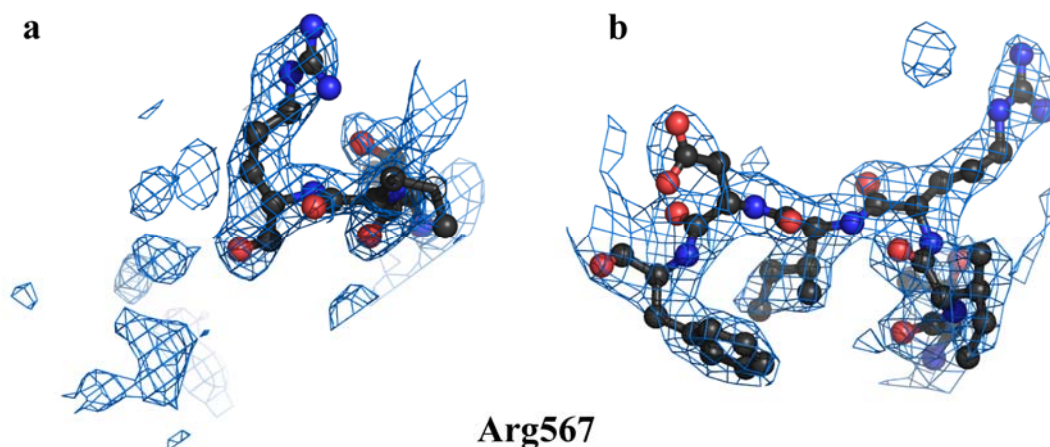


Figure III.24 : Carte de densité électronique d'une région difficilement interprétable. a) Il n'est pas possible de construire la protéine dans cette région alors qu'après le remplacement moléculaire (b) il ne subsiste plus aucun problème.

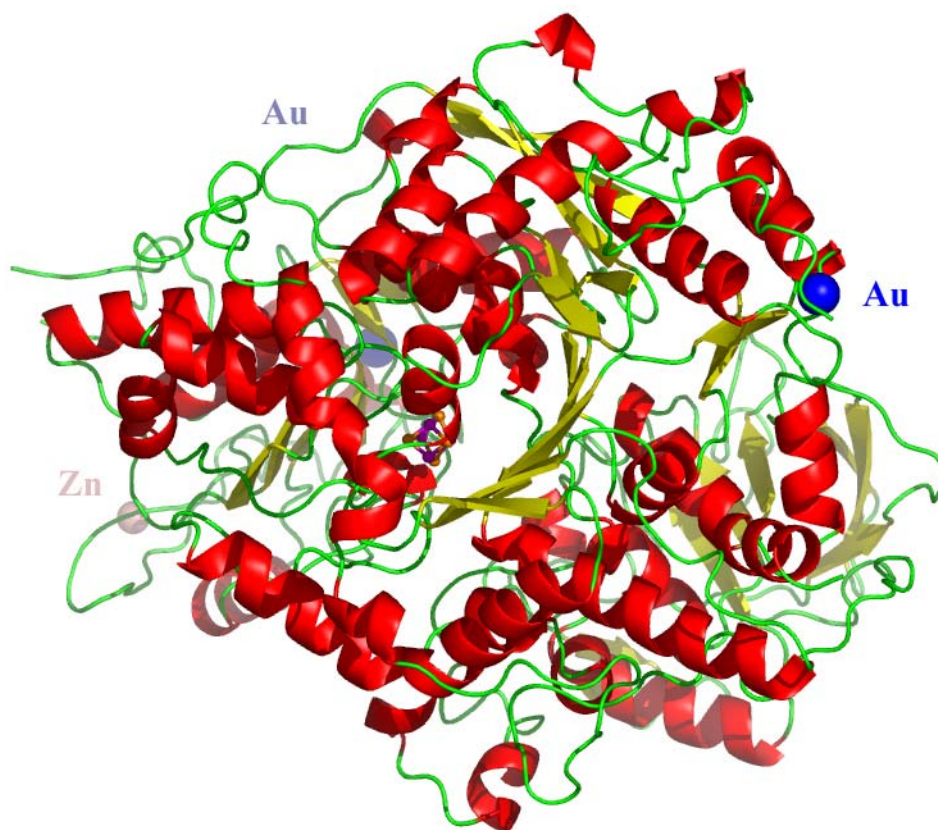


Figure III.25 : Structure finale d'IRP1. Les trois diffuseurs anomaux ont été positionnés sur la structure.

Le modèle définitif a pour facteur R 18% et pour R_{libre} 22,5% pour une résolution de 2,5 Å (Figure III.25).

3.4 Affinement

Toutes les étapes d'affinement ont été réalisées avec le programme REFMAC (Murshudov et al., 1997). 5% des réflexions ont été mis de côté afin de calculer le R_{libre} . Nous n'avons pas modélisé le solvant à basse résolution jusqu'à ce que notre modèle soit complet pour cela nous avons appliqué une troncature à basse résolution à partir de 8 Å. Après chaque cycle, des cartes de densité électronique de type 2mFo-DFc et mFo-DFc ont été calculées avec le programme FFT (CCP4, 1994) afin d'effectuer les corrections manuelles. La procédure d'affinement s'est bien déroulée et à chaque cycle, les facteurs R et R_{libre} ont bien diminué. L'évolution du facteur R_{libre} a permis de nous assurer que nous étions sur la bonne voie. Le tableau III.12 rassemble quelques statistiques concernant l'affinement avec le programme REFMAC (Murshudov et al., 1997).

Statistiques d'affinement	
Gamme de résolution (Å)	47 - 2,54
σ cutoff	aucun
Nb de réflexions (utilisées/libres)	27179 / 1435
Nb d'atomes (protéine/autre)	7076 / 501
R / R_{libre}	0,180 / 0,225
RMSD	Liaisons (Å)
	Angles (°)
	Plans (Å)
Facteur B moyen (Å ²)	32,4

Tableau III.12 : Statistiques en fin d'affinement avec le programme REFMAC (Murshudov et al., 1997). RMSD (*root mean square deviation*) : écart-type.

Le modèle obtenu est de très bonne qualité. Le diagramme de Ramachandran correspondant indique que 90% des résidus sont dans les régions les plus favorables et seul la valine 188 et la serine 436 appartiennent à une région défavorable (Figure III.26). L'analyse de la stéréochimie de la protéine par le programme PROCHECK (Laskowsky et al., 1993) indique que le modèle se trouve au dessus de la moyenne des structures affinées à 2,5 Å de résolution.

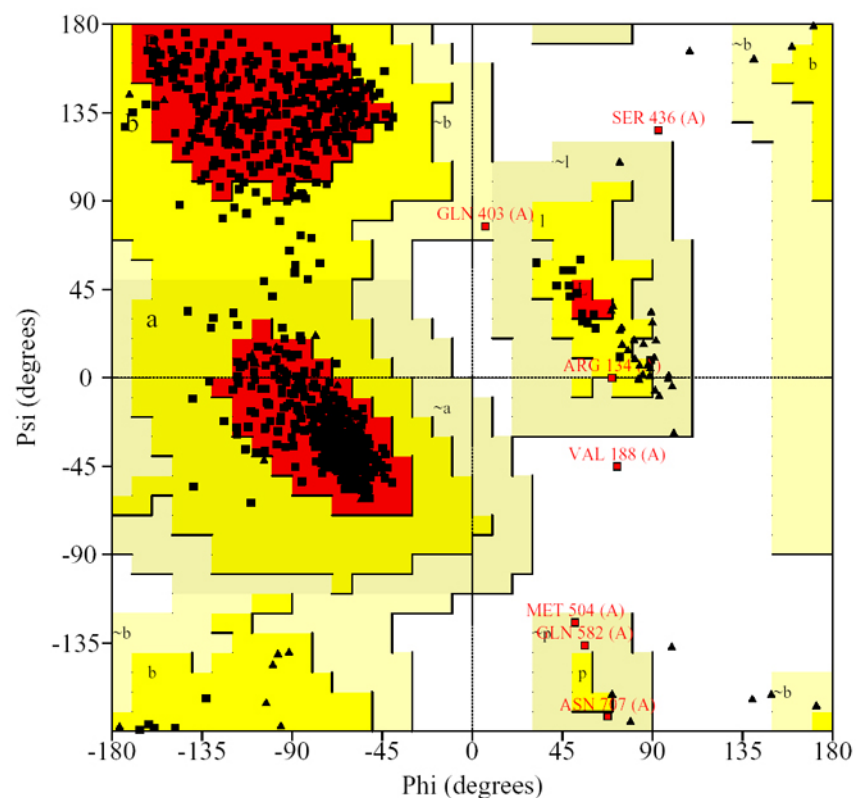
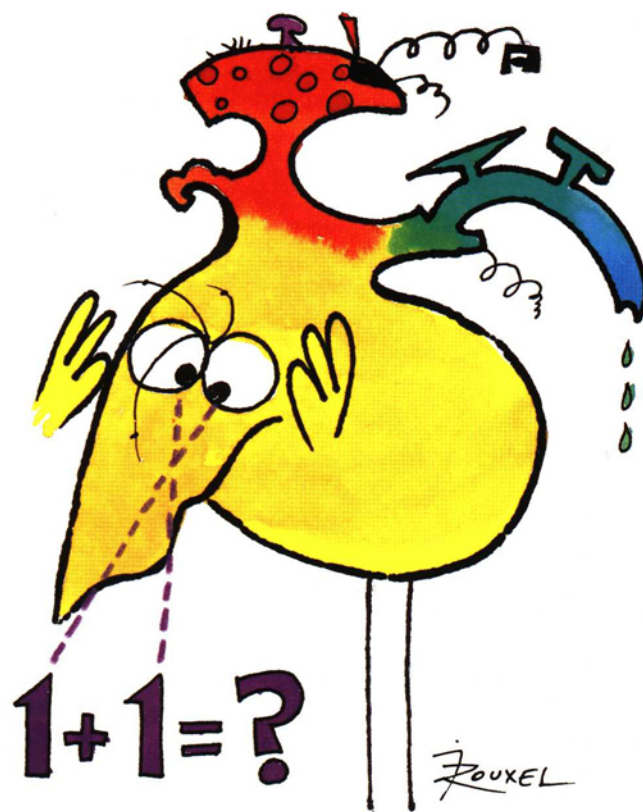


Figure III.26 : Diagramme de Ramachandran, correspondant à la molécule IRP1 de l'unité asymétrique, calculé avec le programme PROCHECK (Laskowski et al., 1993). Les zones en rouges correspondent aux zones autorisées, celles en jaunes aux zones additionnelles autorisées, celles en jaune pâle aux zones généreusement autorisées et celles en blanc aux zones interdites.

Analyse de la structure d'IRP1

Les devises Shadok



LA PLUS GRAVE MALADIE
DU CERVEAU C'EST DE
RÉFLÉCHIR.

Chapitre IV : Analyse de la structure d'IRP1

1 IRP1 versus l'aconitase mitochondriale (mAcn)

1.1 Description générale

La protéine IRP1 est un membre de la famille des aconitases. Les études structurales menées sur cette famille, qui est bien caractérisée biochimiquement, ont permis de mettre en évidence l'existence de pas moins de trois organisations tridimensionnelles différentes pour une seule et même fonction, l'isomérisation réversible du citrate en isocitrate (Figure IV.1).

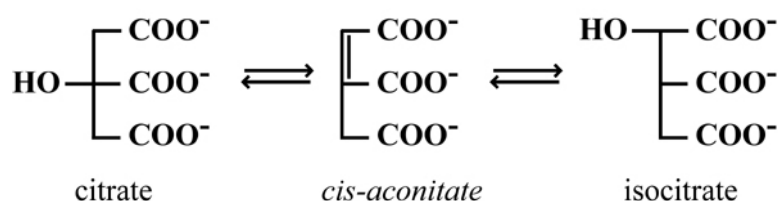


Figure IV.1 : Réaction d'isomérisation du citrate en isocitrate avec son intermédiaire réactionnel, le *cis*-aconitate.

La famille des aconitases se subdivise en trois sous-groupes ; celui dont fait partie l'aconitase mitochondriale qui a un enchaînement continu de ses domaines à savoir que le domaine 1 se trouve du côté amino-terminal et que le domaine 4 est du côté carboxy-terminal, vient ensuite celui des aconitases bactériennes qui est constitué dans sa partie amino-terminale par le domaine 4 auquel se succèdent les domaines 1, 2 et 3 puis enfin un dernier groupe assez hétérogène formé par les isopropylmalate isomérases qui peuvent non seulement adopter une organisation similaire à celle des aconitases mitochondriales mais aussi, chez les bactéries, qui peuvent former des hétérodimères dans lesquels le domaine 4 est une sous-unité à part entière. Néanmoins, ces trois architectures ont toutes en commun l'existence de quatre domaines distincts segmentant les protéines de cette famille (Gruet et al., 1997). IRP1 ne déroge pas à cette règle et la structure de la protéine IRP1 est repliée en quatre domaines. La protéine a des dimensions de $84 \times 79 \times 56 \text{ \AA}^3$ (Figure IV.2 a et b).

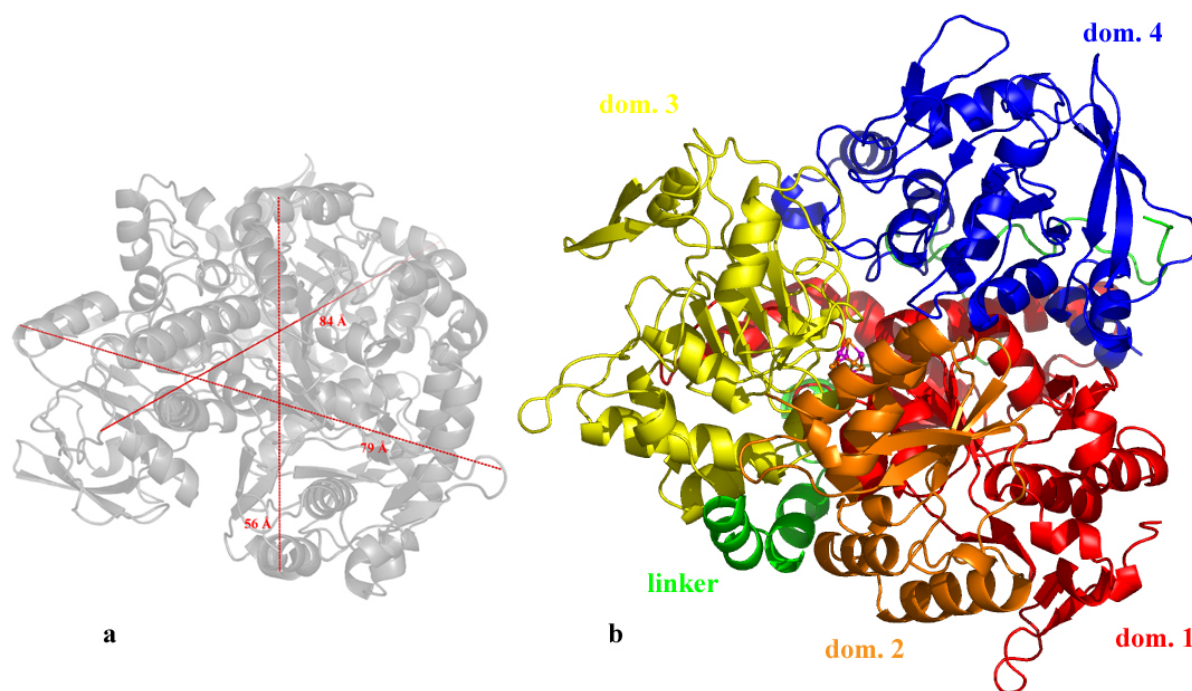


Figure IV.2 : Structure de la protéine IRP1. a) Dimension de la protéine selon trois axes représentants les distances entre Cα les plus longues et b) organisation de la protéine avec une succession caractéristique des domaines, similaire à celle qui est trouvée chez l'aconitase mitochondriale de bœuf. Les domaines 1 (résidu 1 à 240), 2 (241 à 368) et 3 (369 à 592) forment la partie amino-terminale, respectivement rouge, orange et jaune, alors que la partie carboxy-terminale est composée du domaine 4 en bleu (660 à 889), ces deux parties étant reliées par un polypeptide en vert (593 à 659) qui sert de lien au sein de la protéine.

Cette organisation en domaines est tout à fait compatible avec celle de l'aconitase mitochondriale de bœuf, mAcn. L'enchaînement des domaines s'effectue sans permutation ni insertion comme ce peut-être le cas d'AcnB, l'aconitase bifonctionnelle de *E. coli* (Williams et al., 2002) et à l'inverse des isopropylmalate isomérases bactériennes, IRP1 ne possède pas deux sous unités mais un seul et unique corps. Les trois premiers domaines sont fortement associés et forment une dépression peu profonde dans laquelle ils se rejoignent de façon contiguë près du centre de la molécule. L'agrégat fer soufre est lié à trois cystéines du troisième domaine à l'intérieur de cette dépression. Le quatrième domaine est rattaché aux trois premiers par un long segment polypeptidique qui sert de lien charnière ou *linker* à la protéine. La forme du quatrième domaine est complémentaire de la surface formée par les trois premiers domaines. L'interface entre le quatrième domaine et les trois premiers est partiellement occupé par des molécules de solvant et comprend majoritairement des contacts entre acides aminés polaires. Néanmoins, certains acides aminés hydrophiles et hydrophobes sont en mesure de faire des contacts directs entre les domaines 1, 2, 3 et le domaine 4. Par conséquent, une fissure étendue est formée conduisant, selon plusieurs chemins d'accès possibles, de la surface à l'agrégat fer soufre au cœur de la protéine. A la fois la charnière et l'hydratation en solvant de la gorge suggèrent la possibilité d'un mouvement d'ouverture qui

pourrait séparer le domaine 4 du reste de la molécule permettant ainsi la diffusion des substrats dans et hors du site actif, de même que la fixation des motifs IRE.

1.2 Analyse par domaine

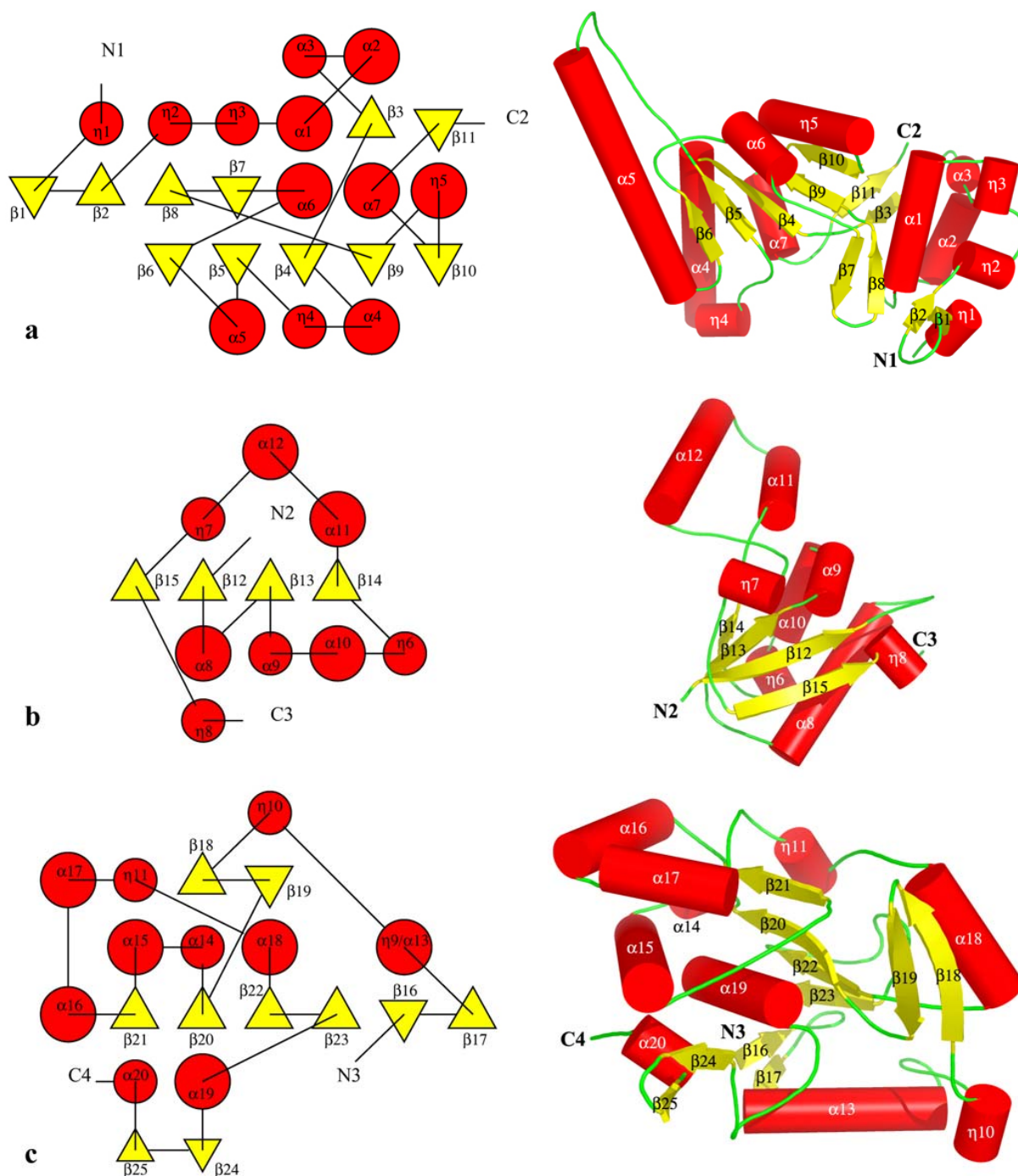


Figure IV.3 : Représentation de l'organisation des éléments de structures secondaires dans les domaines 1 (a), 2 (b) et 3(c). A gauche représentation topologique en deux dimensions et à droite représentation tridimensionnelle. β représente les feuillets β , α les hélices α et η les hélices 3_{10} .

La topologie des éléments de structure secondaire des domaines 1 à 3 contient un motif généralement observé dans les déshydrogénases et d'autres enzymes possédant un

domaine de liaison aux nucléotides (Richardson, 1981). Ces domaines comprennent respectivement cinq, quatre et de nouveau quatre brins β parallèles (Figure IV.3).

De plus les brins β antiparallèles, les hélices α et 3_{10} contribuent à la forte association ternaire de l'ensemble. Le quatrième domaine est, quant à lui, composé d'un tonneau de feuillets β constitué par quatre brins β parallèles et par 3 brins β antiparallèles (Figure IV.4).

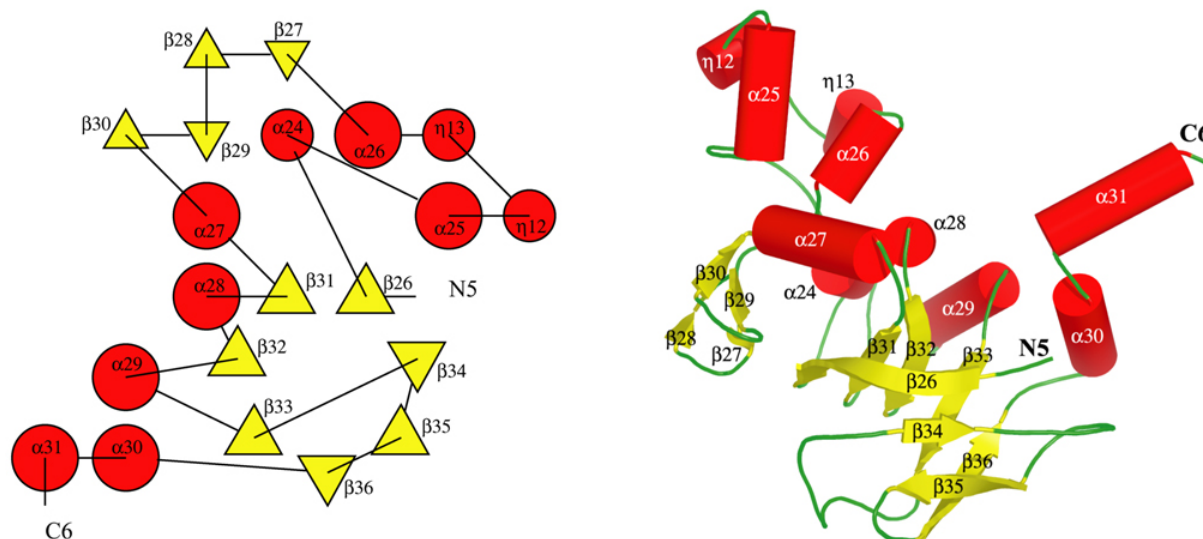


Figure IV.4 : Représentation de l'organisation des éléments de structures secondaires dans le domaine 4. A gauche représentation topologique en deux dimensions et à droite représentation tridimensionnelle.

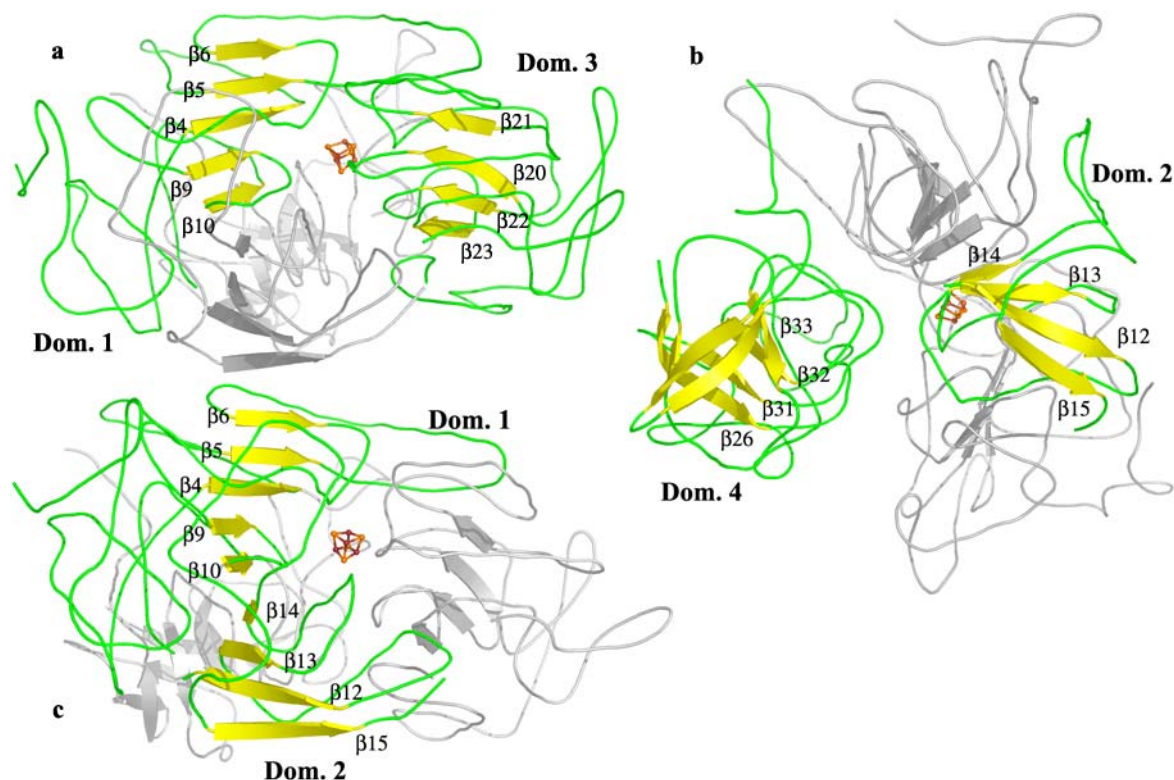


Figure IV.5 : Organisation des feuillets de brins β parallèles. a) comparaison de l'organisation des brins β dans les domaines 1 et 3, b) dans les domaines 2 et 4 et c) dans les domaines 1 et 2.

Les feuillets β parallèles des domaines 1 et 3 ont tous leurs parties carboxy-terminales approximativement orientées vers l'agrégat fer soufre au cœur de la protéine (Figure IV.5 a). En revanche, le domaine 2 a les extrémités amino-terminales de ses brins β de son feuillet parallèle orientés vers la gorge d'accès à l'agrégat alors que l'orientation est inversée pour le domaine 4 (Figure IV.5 b). De plus, les 2 feuillets parallèles des domaines 1 et 2 s'organisent de façons antiparallèles mais sans pour autant former un grand feuillet de brins β du fait même de l'absence de liaison hydrogène entre les brins $\beta 10$ et $\beta 14$ (Figure IV.5 c).

La structure, dans son ensemble, est très proche de celle de l'aconitase mitochondriale (Lauble et al., 1992). Si la structure des domaines de la protéine IRP1 est comparée à celle de mAcn, il est possible de constater que les domaines 2 et 4 sont les mieux conservés et les plus facilement superposables (Tableau IV.1).

IRP1	Nb de résidus	rms (Å)
Domaine 1	133 (239)	2,71
Domaine 2	101 (128)	1,52
Domaine 3	129 (224)	2,75
Domaine 4	159 (230)	1,97

Tableau IV.1 : Tableau récapitulant les écarts types entre les domaines d'IRP1 et ceux de l'aconitase mitochondriale. Les valeurs entre parenthèses représentent le nombre de résidus par domaines.

La région charnière ou *linker* n'est en revanche en rien homologue à celle de l'aconitase. Seul un brin β est présent dans mAcn alors que trois hélices α , $\alpha 21$, $\alpha 22$ et $\alpha 23$ structure le *linker* d'IRP1. Il est intéressant de constater qu'un peu moins d'un cinquième des acides aminés s'ajoutant à la séquence d'IRP1 se fait sur cette seule charnière. En effet, le *linker* est plus long dans IRP1 que celui de l'aconitase laissant supposer une liberté de mouvement encore plus important entre les domaines 1 à 3 et le domaine 4. Ceci semble être en accord avec le fait que, contrairement à mAcn, IRP1 n'a pas pour seule fonction celle d'une isomérase mais aussi celle d'un régulateur post-transcriptionnel. Ceci justifie la nécessité d'une plus grande flexibilité afin de permettre à IRP1, hormis l'accessibilité du citrate au site catalytique, la fixation des séquences nucléotidiques que sont les IRE. Malgré le fait que chaque domaine présente une certaine conservation entre IRP1 et mAcn, IRP1 possède un plus haut degré de complexité dans son organisation tridimensionnelle que l'aconitase mitochondriale. Ainsi, ne serait-ce que pour le domaine 1, pas moins de quatre hélices et autant de brins β viennent s'ajouter à ceux déjà présent dans mAcn. De la même manière un feuillet de deux brins β antiparallèles et quatre hélices s'ajoutent au domaine 3.

Malgré la faible identité reliant IRP1 à l'aconitase mitochondriale, il est intéressant de constater que les deux structures sont relativement proches l'une de l'autre surtout dans l'environnement direct du cœur réactionnel de la protéine (Figure IV.6). Ceci semble important tant il est clair que la protéine IRP1 catalyse la même réaction que son homologue mitochondriale et avec une aussi grande efficacité.

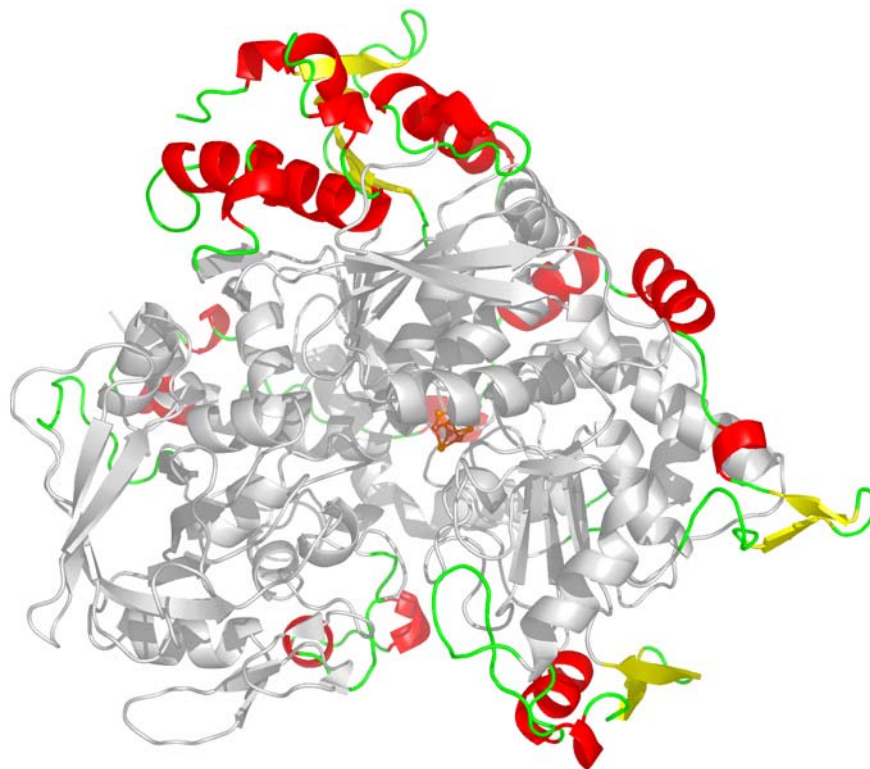


Figure IV.6 : Conservation des éléments de structure de mAcn chez IRP1. Les structures identiques à mAcn chez IRP1 sont colorées en gris alors que le rouge et le jaune servent à modéliser les hélices et les brins β venant s'ajouter à IRP1, ou dont la structure est différente de celle de l'aconitase mitochondriale. Le vert, quant à lui, représente les boucles non structurées.

1.3 L'activité aconitase d'IRP1

L'alignement des séquences d'IRP1, d'IRP2 et de l'aconitase mitochondriale de bœuf n'a pu être réellement efficace qu'après une construction complète de la protéine IRP1 et son alignement structural *in silico* avec l'aconitase (Figure IV.7).

1.3.1 Site actif

Sur les bases de cet alignement de séquence et de la structure tridimensionnelle d'IRP1, nous avons pu nous pencher sur l'activité aconitase de la protéine. Afin de pouvoir jouer ce rôle la protéine doit posséder les acides aminés nécessaires et suffisant à la réaction. Le site actif d'IRP1 est composé de l'agrégat fer soufre et des chaînes latérales des 21 acides aminés, répartis sur les quatre domaines, identifiés comme indispensables à l'activité d'isomérase de la protéine.

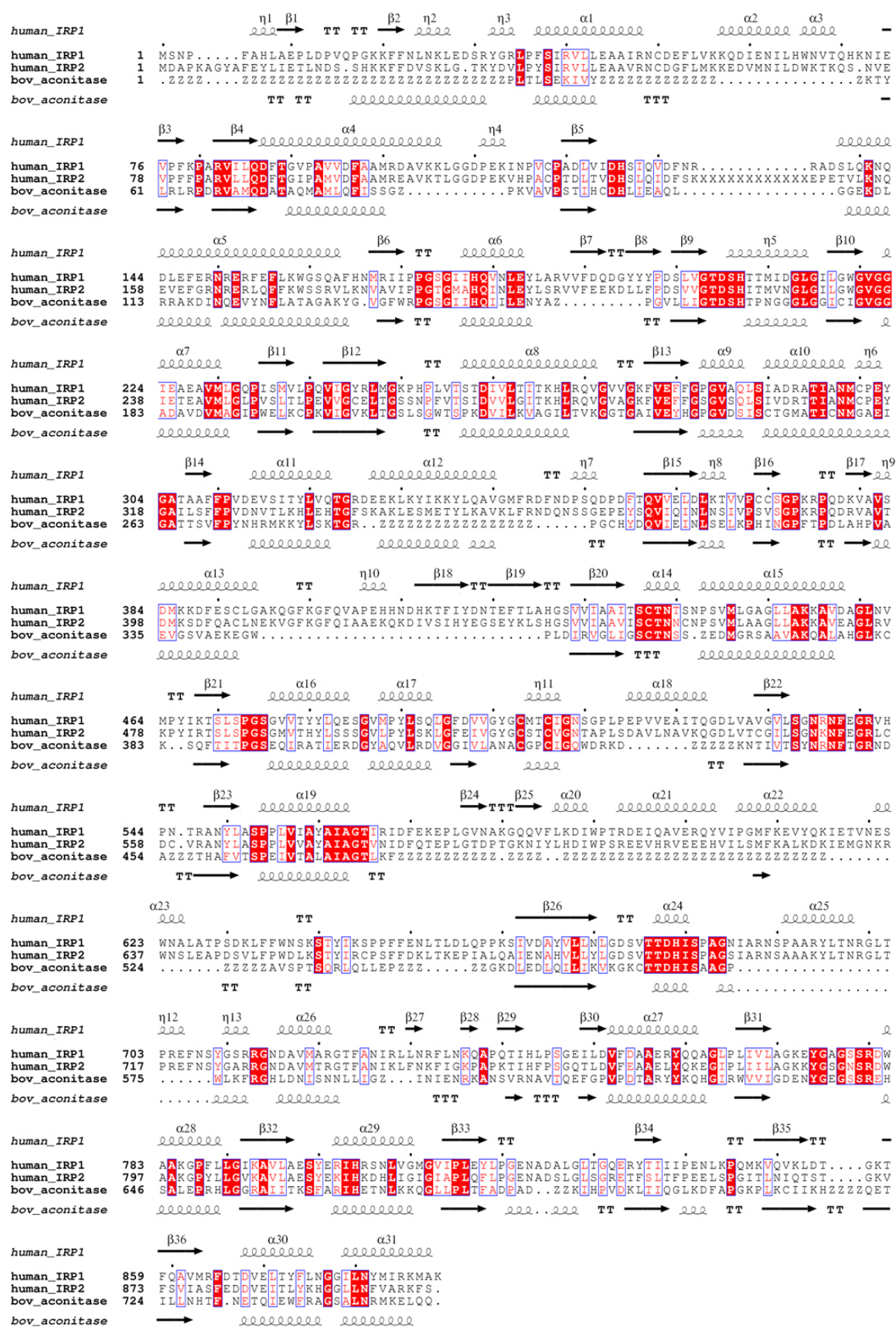


Figure IV.7 : Alignement de séquences entre la protéine IRP1 humaine, IRP2 et l'aconitase mitochondriale. Les XXX correspondent à la zone où les 73 acides aminés supplémentaires d'IRP2 s'insèrent et les ZZZ signalent les portions de l'aconitase mitochondriale qui ne se superposent pas avec IRP1.

Ces chaînes latérales entourent une grande cavité polaire à l'intérieur de la fissure dans laquelle les quatre domaines se retrouvent pour former une dépression peu profonde. Les résidus sont considérés comme faisant partis du site actif s'ils sont en contact avec l'agrégat ou s'ils forment des liaisons avec le substrat ou, enfin, s'il forment des liaisons hydrogène avec un résidu directement en contact avec le substrat (Figure IV.8).

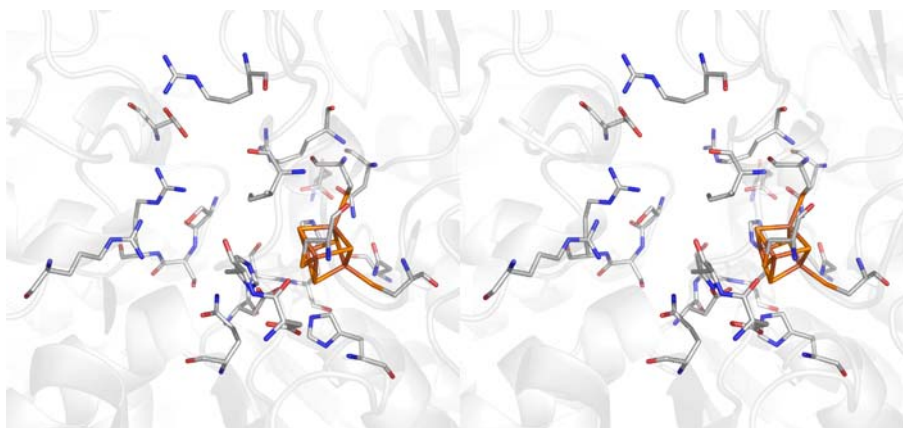


Figure IV.8 : Représentation stéréoscopique du site actif d'IRP1. Ce schéma permet de se rendre compte de l'importance du site actif et de sa complexité.

Le domaine 1 fournit sept résidus au site actif : Gln 86, Asp 125 et His 126 proviennent des boucles connectant les feuillets $\beta 4$ et $\beta 5$ aux hélices $\alpha 4$ et $\alpha 5$; Asp 205, Ser 206 et His 207 forment les premiers tours de l'hélice $3_{10} \eta 5$ et His 178 appartient à l'hélice $\alpha 6$ adjacente. Le domaine 2 ne contribue au site actif qu'avec seulement deux acides aminés Glu 302 de l'hélice $3_{10} \eta 6$ et Asn 298 de l'hélice $\alpha 10$ qui est centrée au milieu du domaine. Le domaine 3 fournit, quant à lui, les trois cystéines, Cys 437, Cys 503 et Cys 506, impliquées dans la coordination de l'agrégat fer soufre et l'isoleucine 507 qui fait un contact hydrophobe avec l'agrégat depuis l'hélice $3_{10} \eta 11$. Le résidu 437 est sur l'hélice $\alpha 14$ alors que les cystéines 503 et 506 sont toutes les deux associées à l'hélice $\eta 11$, au milieu d'une longue boucle reliant les hélices $\alpha 17$ à $\alpha 18$. Une autre boucle, reliant le brin $\beta 22$ au brin $\beta 23$, supporte les acides aminés Asn 535, Arg 536 et Arg 541 eux aussi impliqués dans la formation du site actif. Le domaine 4 apporte les cinq derniers résidus associés à la formation du site actif. L'arginine 780 et les sérines 778 et 779 font parties de la boucle reliant le brin $\beta 31$ à l'hélice $\alpha 28$ alors que Arg 713 est sur la boucle faisant le lien entre $\eta 13$ et $\alpha 26$ et que Asp 678 fait partie intégrante de l'hélice $\alpha 24$.

Les résidus du site actif peuvent être regroupés de la manière suivante : les acides aminés impliqués dans les interactions avec l'agrégat fer soufre sont les cystéines 437, 503 et 506 qui lient trois (Fe1, Fe2 et Fe3) des quatre fers de l'agrégat ensuite l'asparagine 535 qui

fait une liaison hydrogène avec le soufre S1 de l'agrégat alors que l'asparagine 298 en fait une avec le soufre de la cystéine 437 et enfin l'isoleucine 507 qui a un contact hydrophobe significatif avec l'agrégat (Figure IV.9).

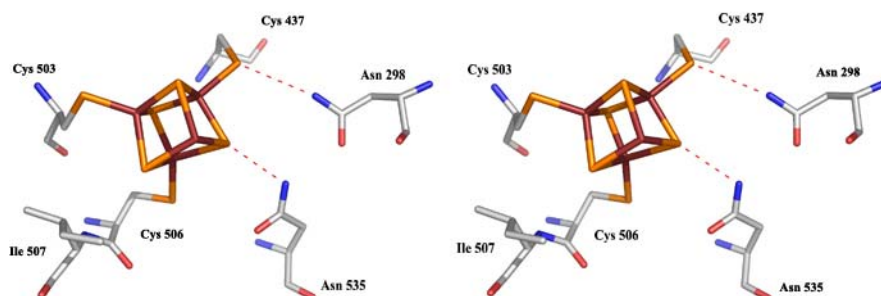


Figure IV.9 : Représentation stéréoscopique des acides aminés impliqués dans des interactions avec l'agrégat fer soufre.

Viennent ensuite les acides aminés interagissant avec le substrat avec en premier lieu les trois paires His - Asp/Glu (Asp 100 - His 101, Asp 205 - His 178 et Glu 302 - His 207), les quatre arginines qui forment des contacts avec les groupements carboxyliques du substrat (les résidus 536, 541, 713 et 780), les résidus qui font des liaisons hydrogène avec le substrat (Gln 86, Ser 206 et Ser 779) et enfin la serine 778 à la base du mécanisme catalytique. Il ne faut pas oublier à cette organisation la présence de l'aspartate 678 qui interagit avec les arginines 541 et 780 via la formation de ponts salins. D'une manière globale, le site actif d'IRP1 est un système complexe de résidus qui forment un réseau de liaisons hydrogène avec le substrat, avec l'agrégat fer soufre, avec les molécules d'eau ainsi qu'avec d'autres résidus de la protéine. L'implication de ces 21 résidus explique, en partie, pourquoi cette enzyme est une aussi grosse protéine. Le site actif présente aussi une charge électrostatique nettement positive à cause de la présence des quatre arginines et du fait que les trois histidines sont appariées avec les aspartates et le glutamate présent dans le site actif (Figure IV.10). Cette caractéristique est clairement constitutive des conditions nécessaires à la fixation de substrat anionique.

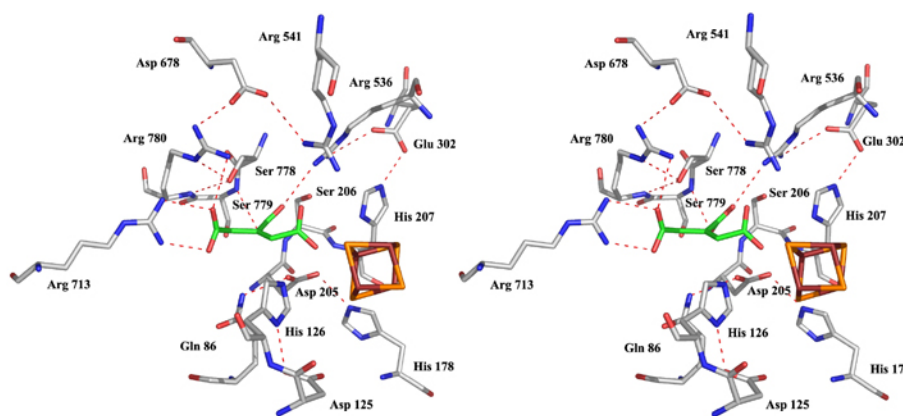


Figure IV.10 : Représentation stéréoscopique des acides aminés impliqués dans des contacts avec le substrat. Dans cette figure, le *cis*-aconitate a été modélisé.

1.3.2 Le mécanisme catalytique

IRP1 catalyse quatre réactions stéréospécifiques : la déshydratation du citrate ou de l'isocitrate en *cis*-aconitate et la réhydratation du *cis*-aconitate en citrate ou en isocitrate. Les différentes études menées sur l'aconitase et IRP1 ont permis de mettre au point un modèle catalytique dans lequel la transformation du citrate en isocitrate se fait via un intermédiaire réactionnel qui est le *cis*-aconitate. Le *cis*-aconitate est censé se lier à l'agrégat fer soufre de deux manières différentes, une appelée mode citrate et l'autre mode isocitrate. Ces deux modes de fixation sont reliés par une rotation de 180° autour de la liaison C α -C β du substrat et l'ensemble de la réaction se produit en trois étapes : la déshydratation du citrate, la rotation du *cis*-aconitate et enfin la réhydratation du *cis*-aconitate pour obtenir l'isocitrate (Figure IV.11).

L'hypothèse actuellement admise est que la complexation du *cis*-aconitate avec l'agrégat modifie l'état de coordination du Fe₄, avec seulement cinq liaisons, suite à la libération, et avant la fixation, d'une molécule d'eau dérivée du groupe hydroxyle partant du substrat. Cette molécule d'eau est facilement échangeable avec le solvant (Beinert et al., 1996). Globalement, ce modèle rend compte de la stéréospécificité de la réaction et de la nécessité d'une *trans* élimination/addition du groupe hydroxyle et du proton à la première et à la troisième étape de la catalyse. L'étape de basculement, qui fait passer le *cis*-aconitate d'un mode de fixation citrate à un mode isocitrate, n'est très probablement pas un processus qui s'effectue en une seule étape. En effet il paraît déraisonnable de penser qu'une seule molécule de *cis*-aconitate est relâchée du Fe₄, réorientée à l'intérieur du site actif et refixée à l'agrégat permettant des interactions plus favorables pour permettre la réhydratation. Il semble plus vraisemblable, aux vues des résultats obtenus sur l'aconitase mitochondriale de bœuf, qu'une molécule de *cis*-aconitate est déplacée par une autre et que cet événement puisse aboutir à la

réorientation du second *cis*-aconitate suivant l'autre mode de fixation. Ce phénomène de déplacement du *cis*-aconitate a été initialement déduit des données de cinétique qui montrait que le relargage du produit était l'étape limitante de la réaction (Schloss et al., 1984). Le déplacement requiert donc l'ouverture et la fermeture de la gorge contenant le site actif, ce qui est en accord avec les changements de conformation observés sur l'aconitase mitochondriale pour la gorge et le *linker* quand la protéine est en présence ou pas de substrat (Beinert et al., 1996). L'ouverture et la fermeture de cette gorge, durant le *turnover*, ne peut se produire que grâce à la présence des chaînes latérales des résidus du site actif occupé qui forment, entre eux et avec les molécules d'eau présentes dans cette région de la protéine, un réseau de liaisons qui empêche la diffusion libre des substrats dans et hors du site actif (Goodsell et al., 1993). Il est à noter cependant, qu'il n'existe aucune évidence directe pour éliminer la possibilité que la même molécule de *cis*-aconitate, formée par la déshydratation du citrate, puisse être réhydratée pour former l'isocitrate.

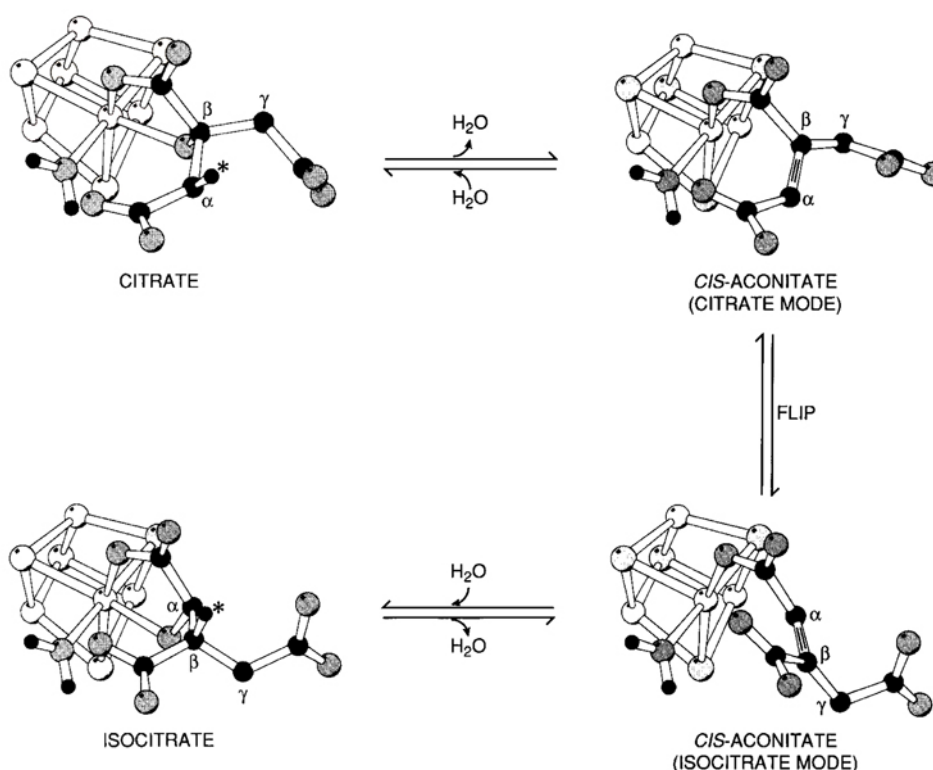


Figure IV.11 : Modèle de mécanisme réactionnel catalysé par IRP1 et l'aconitase. Le produit intermédiaire, le *cis*-aconitate, est proposé comme se fixant de deux manières différentes selon une rotation de 180° par rapport à la double liaison C α -C β . L'astérisque note le proton qui est stéréospécifiquement enlevé au citrate pour être remplacé sur l'isocitrate. L'atome d'hydrogène de la molécule d'eau lie le seul fer non coordonné par les cystéines, Fe4. D'après Lauble et al. (1994)

En s'appuyant sur des études de mutations et de cinétique menées sur l'aconitase, un certain nombre de résidus ont pu être identifiés comme étant critiques à la réaction

d'isomérisation du citrate en isocitrate. L'ensemble de ces résultats a permis de proposer un mécanisme spécifique aux transformations chimiques catalysées par l'aconitase et donc, selon le même modèle, par IRP1 (Zheng et al., 1992 ; Lauble et al., 1992).

La protéine IRP1 libre de tout substrat contient un agrégat $[4\text{Fe-4S}]^{2+}$ avec une liaison entre son Fe4 et un hydroxyle. Après la fixation du citrate l'hydroxyle est protonée et la coordination du Fe4 augmente passant de quatre liaisons à six (Figure IV.12 a). La liaison hydrogène que forme l'histidine 126 avec l'hydroxyle de l'isocitrate suggère que le proton, donné pour former la molécule d'eau, est originaire de cette histidine (Figure IV.12 a). Néanmoins, le proton peut aussi être donné par l'histidine 207 qui est aussi très proche de l'hydroxyle. Ainsi les histidines 126 et 207 sont appariées avec les carboxylates de l'aspartate 125 et du glutamate 302 ce qui les rendent plus propices à la protonation. Le résidu qui arrache le proton du citrate et de l'isocitrate est la sérine 778. En effet, l'oxygène γ de sa chaîne latérale est très proche de l'hydrogène du substrat qui est arraché. De plus, l'environnement de cet atome O_γ , dans le site actif, suggère que son pKa est faible du fait qu'il accepte des liaisons hydrogènes à la fois de l'amide et de la chaîne latérale de l'arginine 780, cette dernière étant positivement chargée (Figure IV.10). De plus des études de mutagenèse dirigées sur cette sérine ont montré que le mutant S642A, l'homologue de la sérine 778 de IRP1 chez l'aconitase, était cinq fois moins actif que la protéine sauvage (Zheng et al., 1992). En outre, des expériences d'échange au tritium ont montré que le proton arraché du substrat était retenu par l'enzyme durant au moins un cycle catalytique, ce qui signifie qu'il est retenu par un groupe qui possède un pKa relativement grand (Rose et O'Connell, 1967). Enfin la géométrie des liaisons hydrogène et de la liaison σ de l'atome O_γ de la sérine est tétraédrique, impliquant que l'alkoxyde de la sérine puisse avoir une paire d'électrons libres orientées vers l'hydrogène du substrat (Figure IV.12 b). En même temps l'hydrogène du C_β de l'isocitrate, qui est en position *trans* par rapport au groupe hydroxyle du C_α , est directement orienté vers l'atome O_γ de la sérine catalytique. La même description du mécanisme réactionnel est valable pour l'hydrogène du C_α du citrate (Figure IV.11).

La fixation du citrate et de l'isocitrate à une forme alkoxyde de l'enzyme est cohérente avec une conservation des charges car le proton tiré de la sérine catalytique est ajouté à l'hydroxyle du Fe4 (Figure IV.12 a). Dans un état protoné la sérine risquerait d'avoir des problèmes d'encombrement stérique avec le citrate ou l'isocitrate ce qui n'est pas le cas du *cis*-aconitate qui a un carbone trigonal à cette même position et ce malgré le mode de fixation qu'il adopte par rapport à l'agrégat. Ceci suggère qu'aucune protonation ou déprotonation de l'hydroxyle ou de la molécule d'eau lié au Fe4 est nécessaire pour que se produise la fixation

ou la libération du *cis*-aconitate ce qui lui facilite son basculement pour les autres étapes de la réactions.

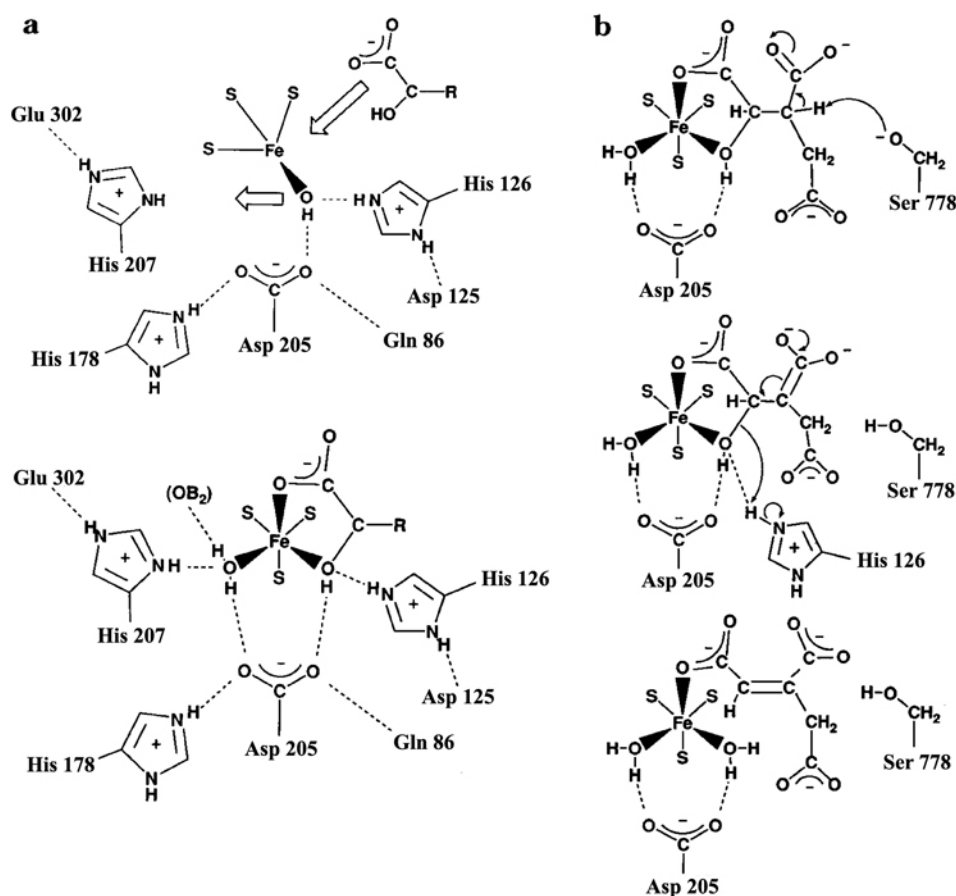


Figure IV.12 : a) Représentation schématique de la transition entre la forme de la protéine dépourvue du substrat et celle l'ayant fixé. Le proton ajouté au Fe-OH pour faire Fe-OH₂ est strictement équivalent au proton dérivé de la sérine 778 qui forme un alkoxyde. Le contact du Fe-OH₂ avec l'oxygène carboxylique du Cβ de l'isocitrate (OB₂) est montré séparément pour plus de clarté. b) Représentation schématique de la réaction transformant l'isocitrate en *cis*-aconitate en supposant bien que la sérine 778 est déprotonnée quand le substrat est fixé et que Fe-OH₂ s'est formé. D'après Beinert et al. (1996).

Faisant suite au départ du proton de Cβ de l'isocitrate un carbanion intermédiaire est formé (Figure IV.12 b). La disparition de cet intermédiaire est concomitante avec la protonation de l'hydroxyle du substrat par l'histidine 126 qui clive la liaison carbone oxygène afin de donner les produits que sont le *cis*-aconitate et l'eau (Figure IV.12 b). La charge négative de l'oxygène carboxylique du Cβ de l'intermédiaire peut être stabilisée par la formation d'une liaison hydrogène avec la molécule d'eau liée au Fe4 selon une interaction de type « low-barrier hydrogen bond ». Cette interaction hydrogène est particulière car elle se produit entre un donneur et un accepteur qui sont très proches l'un de l'autre et dont les pKa sont quasiment identiques. Sous ces conditions le proton est également réparti entre le donneur et l'accepteur conduisant à une réduction significative de l'ensemble de l'énergie

libre du système (Cleland, 2000). Dans l'aconitase, il est possible que la molécule d'eau liée au Fe4 ait un pKa approchant celui de l'intermédiaire acide à cause de la coordination du métal. De plus, la géométrie tétraédrique suggère que la molécule d'eau accepte de faire une liaison hydrogène avec l'histidine 207 tout en faisant une liaison avec l'aspartate 205. En conséquence de quoi, un des hydrogènes de la molécule d'eau est orienté vers le groupe carboxylique du C β de l'isocitrate. Toutes ces données confirment que l'alkoxyde de la sérine est à la base de la réactivité de l'aconitase mitochondriale, et donc par homologie de structure et identité de séquence à la base de l'activité aconitase d'IRP1, et que la « *low-barrier hydrogen bond* » stabilise le carbanion intermédiaire de l'état de transition.

1.3.3 Tunnel d'accessibilité

Au pH physiologique, le citrate a ses fonctions carboxyliques complètement déprotonées. Comme il a été décrit plus haut la cavité qui accueille le substrat présente une charge globale largement positive ce qui se comprend par les propriétés anioniques du citrate et de l'isocitrate. Il existe de ce fait, non pas un, mais deux chemins d'accès qui permettent au substrat de rentrer dans la protéine (Figure IV.13). Ces deux chemins sont bordés par plusieurs acides aminés chargés positivement. Il s'agit des arginines 269, 375, 536, 541, 547, 694 et 780 ainsi que des histidines 207 et 679. Sur l'ensemble de ces acides aminés trois sont d'ailleurs impliqués dans la constitution du site actif, Arg 780, Arg 536 et His 679. Il va de soi que les mouvements du domaine 4 par rapport au reste de la protéine peuvent permettre une plus grande plasticité au niveau de ces chemins d'accès.

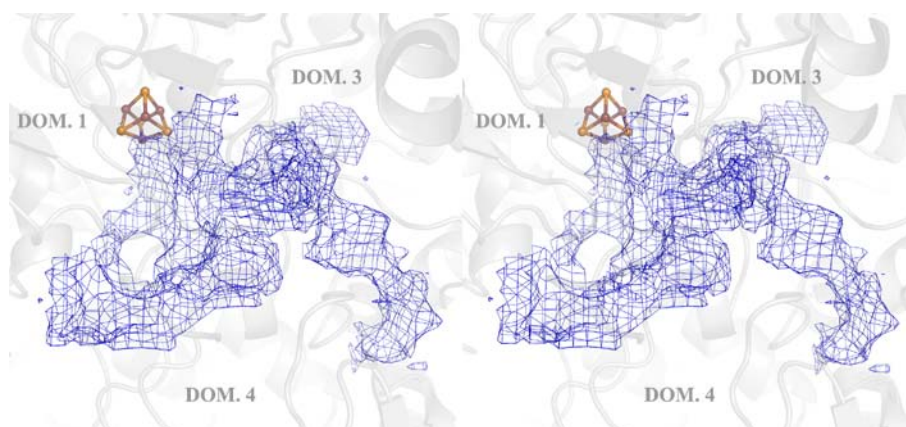


Figure IV.13 : Représentation des deux chemins d'accès à l'agrégat depuis la surface de la protéine. La carte représentant cette accessibilité a été réalisée avec le programme CAVENV (Volbeda, 1999) et selon une sonde avec un diamètre de 1 Å.

L'aconitase, contrairement à IRP1, ne présente pas autant de chemin d'accès à l'agrégat. Il est possible de penser que la plus grande flexibilité, et donc mobilité, du domaine

4 d'IRP1 permette cette multiplication des tunnels d'accessibilité. Ceci semble être en accord avec le fait qu'IRP1 n'a pas pour seule fonction celle de catalyser l'isomérisation du citrate en isocitrate mais aussi celle de fixer les motifs IRE contenus dans certains ARN messagers.

1.3.4 Le site zinc

La différence majeure qui existe entre les premières formes cristallines, P2₁ et C2, et la nouvelle forme orthorhombique C222₁ est sûrement due à la présence de cet atome de zinc dans la préparation des derniers échantillons. En effet cet atome de zinc, identifié par fluorescence X (voir § III.2.3.1), participe à l'empilement cristallin car il se situe entre deux molécules IRP1 reliées par une symétrie cristallographique d'ordre 2 (Figure IV.14).

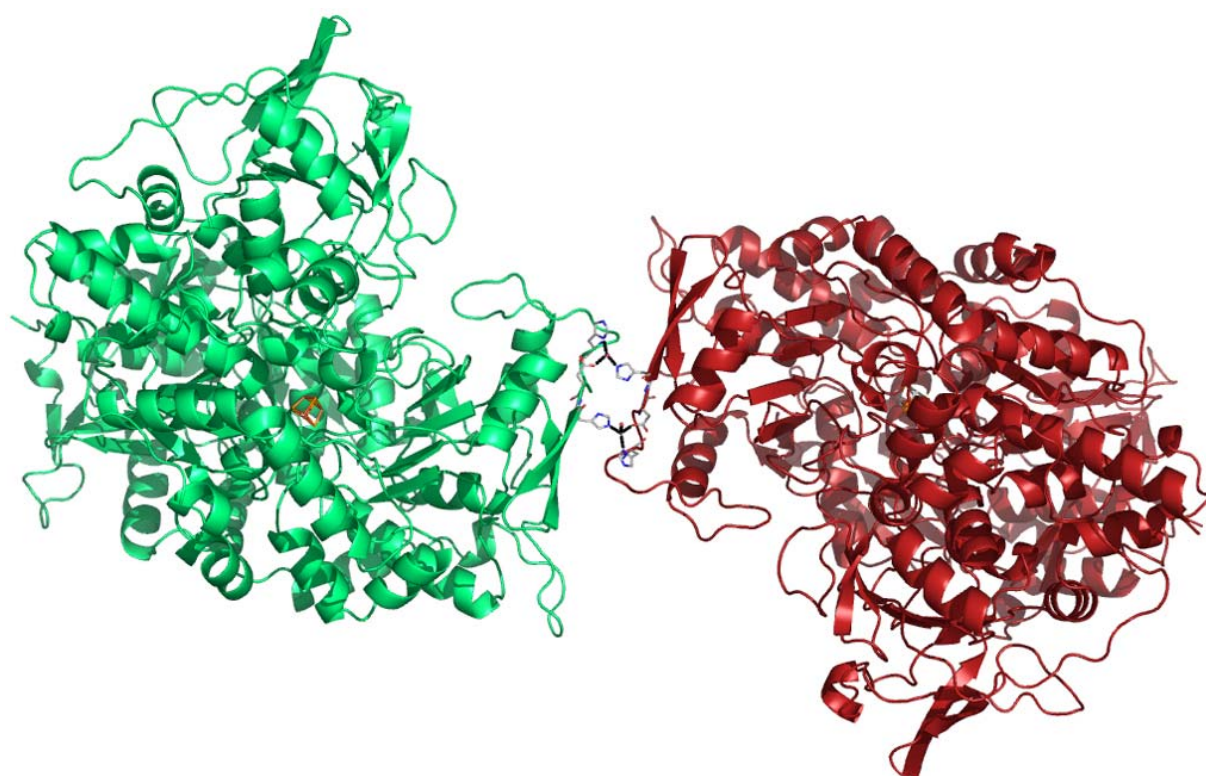


Figure IV.14 : Vue des deux molécules symétriques mettant en commun les acides aminés nécessaires à la coordination des atomes de zinc.

Chaque molécule apporte un atome de zinc ainsi qu'une partie des acides aminés nécessaires à leur coordination tétraédrique. Ainsi lorsque la première molécule apporte le zinc elle fournit aussi une histidine pour sa coordination, His 412, et une molécule d'eau alors que son symétrique apporte un aspartate, Asp 411, et une histidine, His 408. L'ensemble permet une bonne coordination du zinc (Figure IV.15).

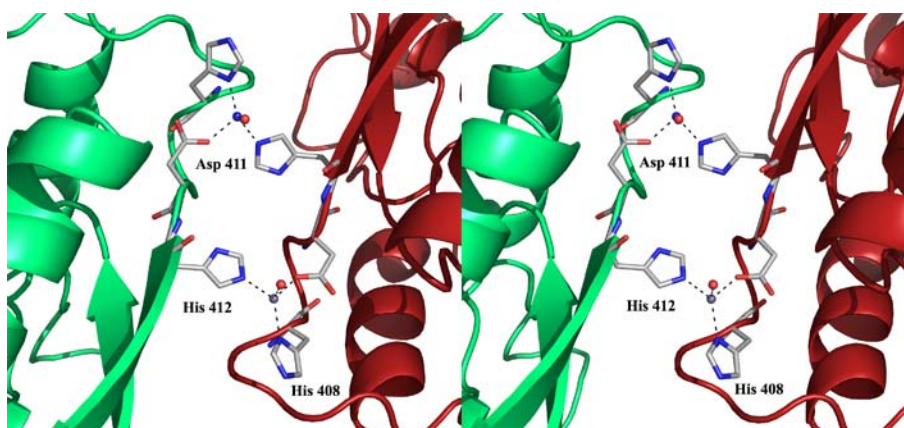


Figure IV.15 : Sites d'interactions entre les molécules symétriques. Chaque atome de zinc est coordonné par deux histidines, un aspartate et une molécule d'eau. Tous ces acides aminés sont partagés entre les deux symétriques.

1.3.5 Sérines phosphorylables

IRP1 est capable de subir une régulation par phosphorylation encore mal définie dans son origine mais qui en revanche a été très bien caractérisée quant à ses cibles sur la protéine. Ainsi deux sérines ont pu être clairement identifiées comme étant des sites potentiels de phosphorylation par la protéine kinase C. Il s'agit de la sérine 138 qui est située en surface de la protéine sur une hélice α et de la sérine 711 qui est sur une hélice 3_{10} un peu plus enfouie (Figure IV.16). La phosphorylation de ces deux sérines entraîne des réponses différentes selon que ce soit la sérine 138 ou la sérine 711 qui en est la cible.

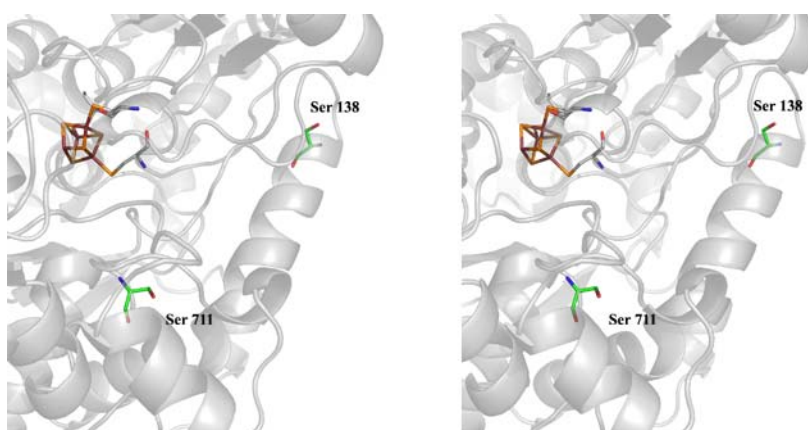


Figure IV.16 : Localisation des deux sérines phosphorylables sur la protéine IRP1. La sérine 138 est exposée au solvant alors que la sérine 711 présente une accessibilité plus réduite.

Dans le cas de la sérine 138 la phosphorylation a pour effet de déstabiliser l'agrégat fer soufre ce qui conduit à son démantèlement et à l'activation d'IRP1. Cette activation consiste, pour la protéine, en l'acquisition de son activité de régulateur post-transcriptionnel des ARN messagers contenant des motifs IRE. Cela permet de moduler, de façon indépendante de la concentration en fer, l'activité de la protéine IRP1. Deux arginines, Arg 134 et 135, une

lysine, Lys 141 et un aspartate, Asp 137, sont dans l'environnement direct de la sérine 138 (Figure IV.17). La phosphorylation de ce résidu entraîne une perturbation de la distribution des charges dans cette région.

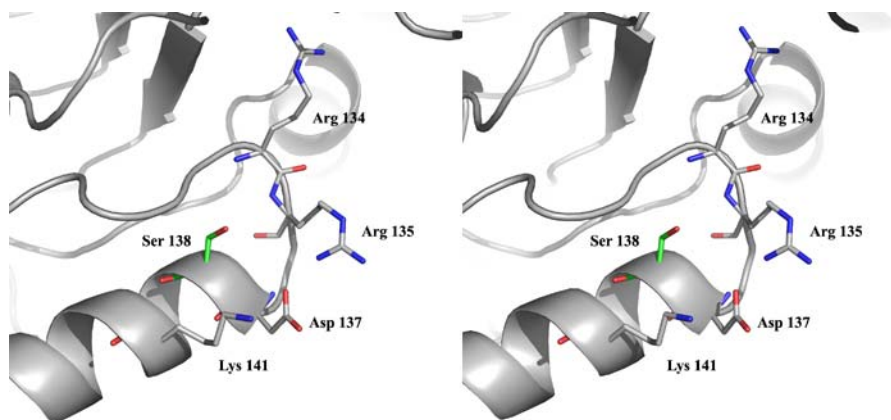


Figure IV.17 : Environnement de la sérine 138.

La sérine 138 est localisée sur l'hélice $\alpha 5$ juste après le brin $\beta 5$. La boucle reliant ces deux éléments de structure secondaire portent deux acides aminés essentiels à l'activité catalytique d'IRP1, l'aspartate 125 et l'histidine 126 (Figure IV.18). De plus cette même boucle jouxte l'hélice 3_{10} sur laquelle les cystéines 503 et 506 sont localisées, cystéines qui coordonnent l'agrégat fer soufre (Figure IV.18). Il est aussi à noter que l'isoleucine 507, qui forme un contact hydrophobe avec l'agrégat, est aussi à proximité de cette boucle (Figure IV.18).

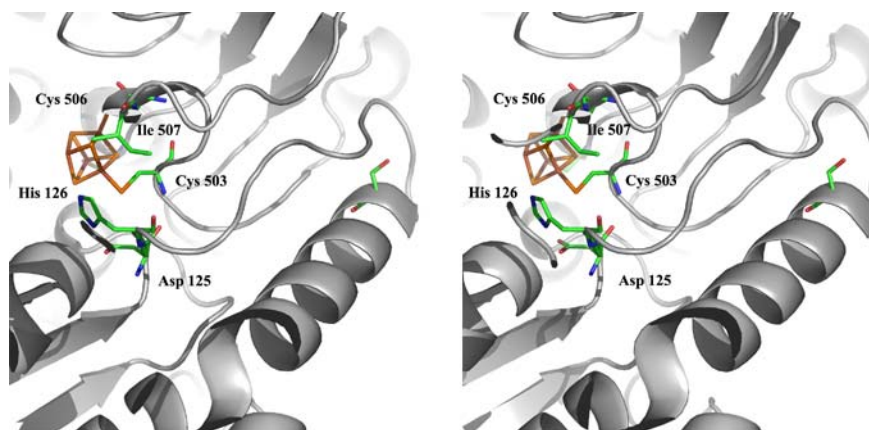


Figure IV.18 : Représentation des acides aminés indispensables au site actif et qui sont aussi à proximité de la sérine 138.

Il est donc possible d'envisager que la phosphorylation de la sérine 138 aboutit à un remaniement complet de cette région de la protéine. La charge négative du phosphate devrait induire une réorganisation des résidus chargés dans son environnement et peut-être du segment d'hélice α portant la sérine 138 ce qui a forcément des répercussions sur ses voisins

directs et donc sur l'agrégat. Il est tout à fait logique d'envisager que les cystéines 503 et 506 ne se retrouvent plus dans une conformation idéale à la coordination de l'agrégat et de ce fait il en est déstabilisé et plus enclin à être désassemblé.

En ce qui concerne la sérine 711, des études biologiques sur la forme aconitase d'IRP1, ont montré que sa phosphorylation bloque la transformation du citrate en *cis*-aconitate mais n'empêche nullement le passage du *cis*-aconitate en isocitrate (Fillebeen et al., 2005). Il a été proposé que cette pseudo-fonction aconitase aurait pour conséquence de promouvoir soit l'accumulation de citrate dans la cellule soit l'élimination de l'isocitrate du milieu et donc, d'une manière détournée, de diminuer la disponibilité en NADPH dans le cytoplasme (Pitula et al., 2004). Contrairement à la sérine 138, l'environnement direct de la sérine 711 ne laisse pas entrevoir de grands changements quant à la distribution des charges dans cette région. En effet, le seul acide aminé basique qui pourrait être perturbé par la présence d'un phosphate n'est autre que l'arginine 713 qui est directement impliquée dans le maintien du substrat dans le site actif (Figure IV.19). Il est donc facilement envisageable d'imaginer que la phosphorylation de cette sérine, légèrement enfouie, pourrait avoir plusieurs effets sur son environnement direct. Il n'est pas impossible de penser que l'ajout d'un phosphate pourrait avoir comme conséquence une modification de l'organisation tridimensionnelle de cette région et ce pour deux raisons : pour commencer la présence d'une charge négative pourrait bouleverser l'arginine 713 qui ne doit pas pouvoir jouer son rôle dans la stabilisation du substrat dans le site actif de la protéine. Dans un deuxième temps l'encombrement stérique du phosphate doit sûrement avoir une répercussion sur les structures secondaires qui l'entourent de même que sa charge doit influencer sur les amines de la chaîne principale.

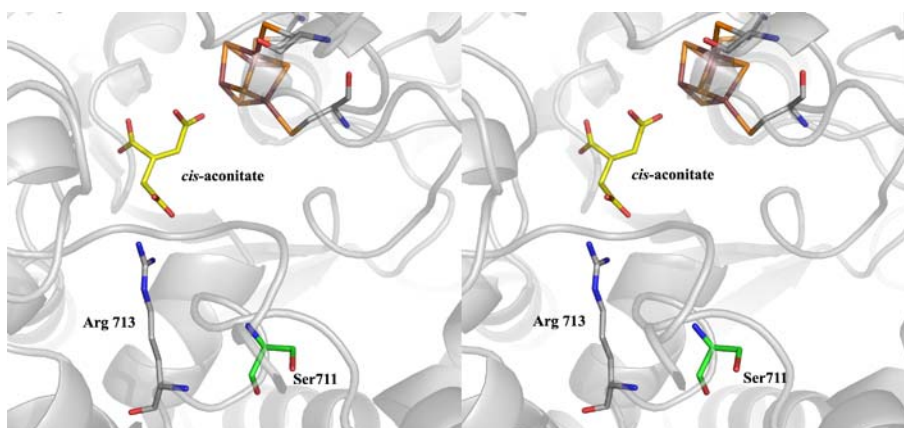


Figure IV.19 : Localisation de l'arginine 713 par rapport à la sérine 711. Le *cis*-aconitate a été modélisé ici en jaune.

En conclusion de quoi, la structure d'IRP1 nous a permis d'avaliser certains résultats obtenus par le passé par des expériences biochimiques et enzymologiques. Ainsi, il est possible, à la vue de la structure d'IRP1, de dire que l'activité aconitase de la protéine correspond trait pour trait à celle qui a été décrite pour celle de l'aconitase mitochondriale. Ceci est en accord avec la parfaite conservation du cœur catalytique de la protéine et surtout de la parfaite identité des acides aminés impliqués dans la réaction d'isomérisation du citrate en isocitrate. De plus, les données à plus haute résolution de la forme P2₁ n'ont pas permis d'avoir une meilleure définition du mode de fixation du substrat. Il semble, en effet, qu'un mélange de conformations ou qu'une faible occupation du site de fixation rendent la densité électronique de cette région non interprétable. La structure d'IRP1 a aussi permis d'envisager les possibles effets de la phosphorylation des sérines 138 et 711. Il semble clair qu'à la vue de leur localisation respective, une modification dans la distribution des charges de ces régions proches de l'agrégat entraînerait forcément une perturbation de l'activité catalytique d'IRP1. Cette perturbation pourrait se traduire, soit pas un démantèlement de l'agrégat, soit par un blocage de l'activité catalytique de la protéine.

2 Modélisation de la fixation du motif IRE

A partir de la structure d'IRP1, et connaissant celle du motif IRE régulant la traduction de la ferritine (Addess et al., 1997), nous avons essayé de réaliser la modélisation de la fixation de l'IRE sur la protéine IRP1. Afin de restreindre le nombre de possibilités quelques règles simples ont été adoptées :

- La littérature scientifique a permis de cribler aussi bien les acides aminés impliqués dans la fixation de l'ARN, que les zones de l'ARN interagissant avec la protéine ;
- la topologie de la surface de la protéine et la répartition des charges sur cette même surface ne permettent pas une libre orientation de l'ARN ;
- la modélisation restant spéculative, un minimum de changements, au sein des structures secondaires du domaine 4 censé bouger pour permettre la fixation de l'IRE, a été effectué ;
- enfin les travaux réalisés par le passé sur l'aconitase mitochondriale de bœuf, servant de modèle à une modélisation des contacts IRE-protéine, ont servi de point de départ au présent travail.

Les résultats de ces dix dernières années ont permis de mettre en évidence l'existence de zones nécessaires et indispensables à la fixation des motifs IRE sur la protéine IRP1. Les

premiers travaux réalisés par mutagenèse sur ce sujet par Philpott et al. (1994) ont abouti à la mise en lumière de l'importance de trois résidus arginines, Arg 536, 541 et 780, dans les interactions avec l'ARN. Ces mêmes travaux ont apporté la preuve que les cystéines, Cys 437, 503 et 506, qui coordonnent l'agrégat fer soufre sont aussi impliquées dans les contacts avec l'IRE. La même année, Basilion et al. (1994) ont réussi à identifier tout un enchaînement d'acides aminés capital à l'interaction. Cette séquence, identifiée par réticulation après une irradiation aux rayons ultraviolets, couvre pas moins de neuf résidus de Asp 121 à Val 130 avec une petite note particulière pour la sérine 127 qui a été retrouvée comme étant fixée de façon covalente à l'ARN suite à cette expérience. En 1996, Butt et al. (1996) confirmèrent par mutagenèse l'importance des arginines 536, 541 et 780. Schlegl et al. (1997) ont réussi, grâce à plusieurs RNases et autres sondes chimiques, à localiser de façon relativement précise les zones de l'IRE protégées lors de son interaction avec la protéine IRP1. Il en résulte que les bases A-G-U de la boucle sont clairement inaccessibles à une dégradation et il en est de même pour la base C non appariée. Les auteurs ont aussi cartographié les riboses qui étaient protégés par les interactions ARN-protéine (Figure IV.20).

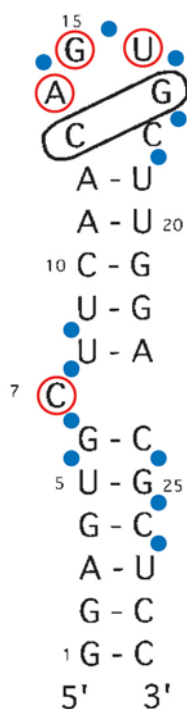


Figure IV.20 : Structure secondaire du motif IRE de l'ARN messager de la ferritine. Les bases qui sont cerclées de rouge sont protégées d'une dégradation lors de l'interaction IRE-IRP1. Les points bleus représentent les riboses qui sont, eux aussi, protégés par cette même interaction. Dans la structure RMN du motif IRE la cytosine 13 forme une liaison avec la guanine 17, ce qui est cerclé de noir sur le schéma.

Arrive ensuite les deux plus importants travaux réalisés sur le sujet. Kaldy et al. (1999), toujours par mutagenèse mais aussi bien sur la protéine que sur le motif IRE, ont réussi à isoler deux segments de protéine impliqués dans la fixation de l'ARN, Asn 685 à Asn 689 et Arg 732 à Gln 737. Ils ont été capables de surcroît d'identifier que la région 685-689

interagirait avec la boucle CAGUGN de l'ARN alors que la portion 732-735 fixerait plutôt la base non appariée de ce même motif, à savoir la cytosine 7. Gegout et al. (1999) ont démontré par protéolyse en présence d'IRE que plusieurs acides aminés, localisables sur deux zones de la protéine, étaient protégés. Il s'agissait des arginines 149 et 187, de l'aspartate 125 et du glutamate 155 tous présents dans le domaine 1 et dans le domaine 4 les arginines 721, 728 et 732 puis l'aspartate 751. Ces résultats apportaient aussi quelques observations intéressantes concernant de nouveaux candidats potentiellement impliqués dans la fixation du motif IRE : Lys 79, Asp 87, Arg 101, Asp 102, Lys 105, Lys 106, Asp 137 et Lys 141. Une mention spéciale pour l'aspartate 137 et la lysine 141 car il semble que la fixation de l'ARN à la protéine entraîne non pas une protection de ces deux acides aminés mais au contraire une surexposition aux protéases.

Fort de ces résultats et des travaux de Paraskeva et Hentze (1996), qui ont été les premiers à proposer « l'accostage » du motif IRE sur l'aconitase mitochondriale, la modélisation a été initiée par l'identification des surfaces de contact entre le domaine 4 et les domaines 1,2 et 3 et par la localisation de tous les acides aminés décrits précédemment (Figure IV.21 a et b).

La première constatation flagrante, qui peut être déduite de cette localisation, est que tous les acides aminés décrits comme étant protégés par la fixation du motif IRE ne le sont pas obligatoirement suite à une interaction directe avec l'ARN. Plusieurs d'entre eux, Lys 79, Asp 121 et Arg 187, sont positionnés sur la face opposée à la fixation de l'IRE (Figure IV.21 b) ce qui laisse à penser que leur protection n'est due qu'à une interaction indirecte de l'ARN. Cela peut être le fruit d'une réorganisation des parties les plus flexibles de la protéine et pour cela le *linker* pourrait être le candidat idéal. Il pourrait facilement subir, au cours de l'ouverture du domaine 4, un déplacement à la surface de la protéine IRP1 et se retrouver dans cette zone protégeant, par sa seule présence, les acides aminés localisés dans cette région. De la même manière il paraît tout à fait clair que la présence de l'agrégat fer-soufre est un frein majeur à la fixation de l'IRE. En effet, la protéine IRP1 dans sa forme IRE, c'est-à-dire avec les capacités de fixer de l'ARN, a besoin d'une portion d'acides aminés, Asp 121 à Val 130, qui se trouve juste à proximité de l'agrégat (Figure IV.21 a).

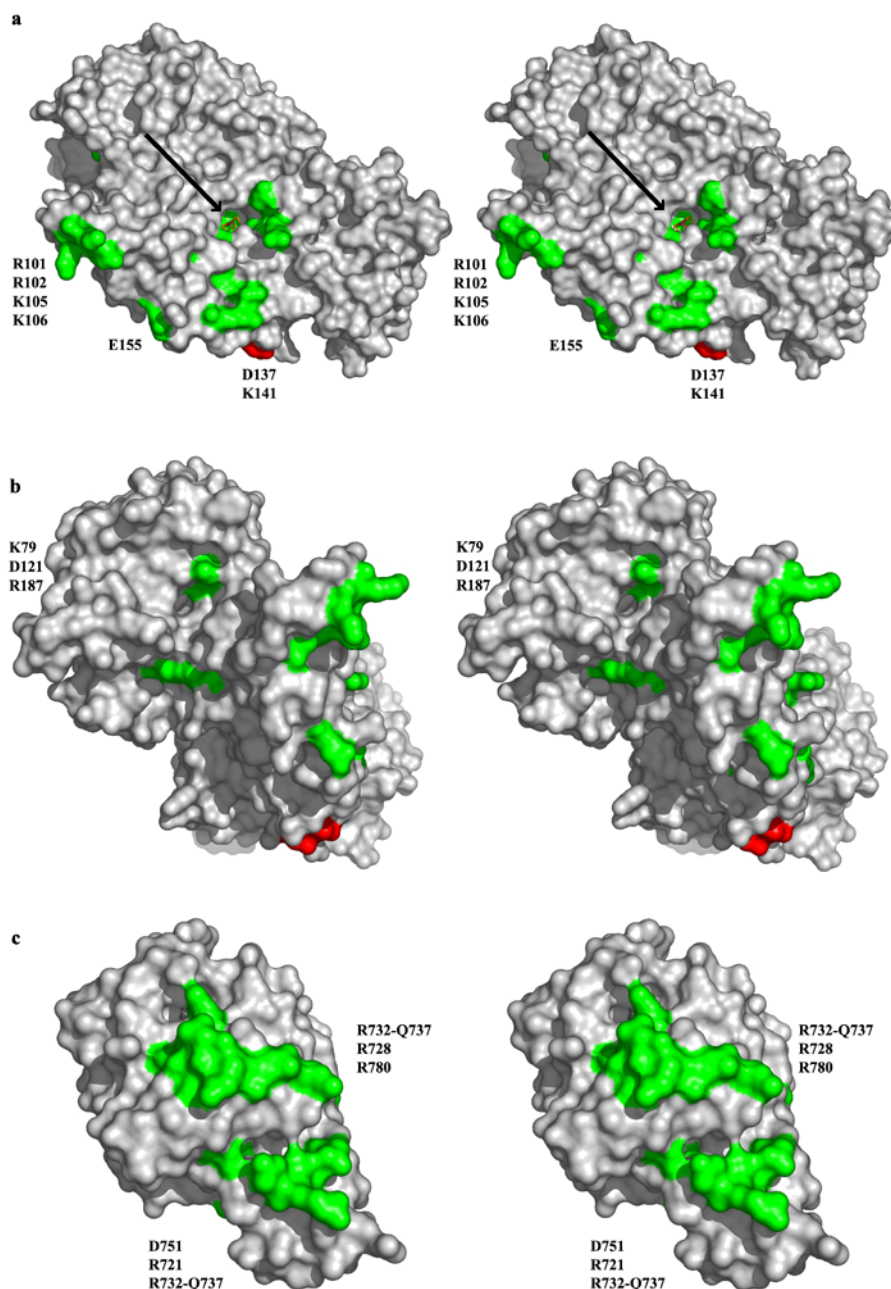


Figure IV.21 : Représentation de la localisation des acides aminés impliqués dans les interactions avec l'IRE. En vert les acides aminés qui sont protégés alors que le rouge symbolise ceux qui sont surexposés aux protéases après la fixation du motif IRE. a) schématise la surface des domaines 1, 2 et 3 vue de face, b) représente la surface des mêmes domaines mais après une rotation de 90° selon l'axe y et c) illustre la surface du domaine 4. La flèche noire indique la position de l'agrégat et seuls, les acides aminés qui seront discutés ultérieurement sont labellisés.

Cela explique en partie la stricte exclusivité de la protéine vis-à-vis de ses deux fonctions. D'une manière plus générale, le positionnement de tous les acides aminés appartenant aux domaines 1, 2 et 3, décrit comme étant important dans la fixation des IRE, tapisse une partie considérable de la surface qui normalement est en contact avec le domaine 4 (Figure IV.22 a). Certain d'entre eux, néanmoins, Arg 101, Asp 102, Lys 105, Lys 106 et Glu

155, sont positionnés de telle façon que seul un remaniement structural et une réorganisation du domaine 4 par rapport aux domaines 1, 2 et 3 peut permettre leur protection (Figure IV.21 b). De la même façon le domaine 4 présente aussi une répartition de ses acides aminés, impliqués dans la fixation de l'ARN, à la surface de la zone complémentaire aux domaines 1, 2 et 3 mais d'une façon moins prononcée (Figure IV.22 b).

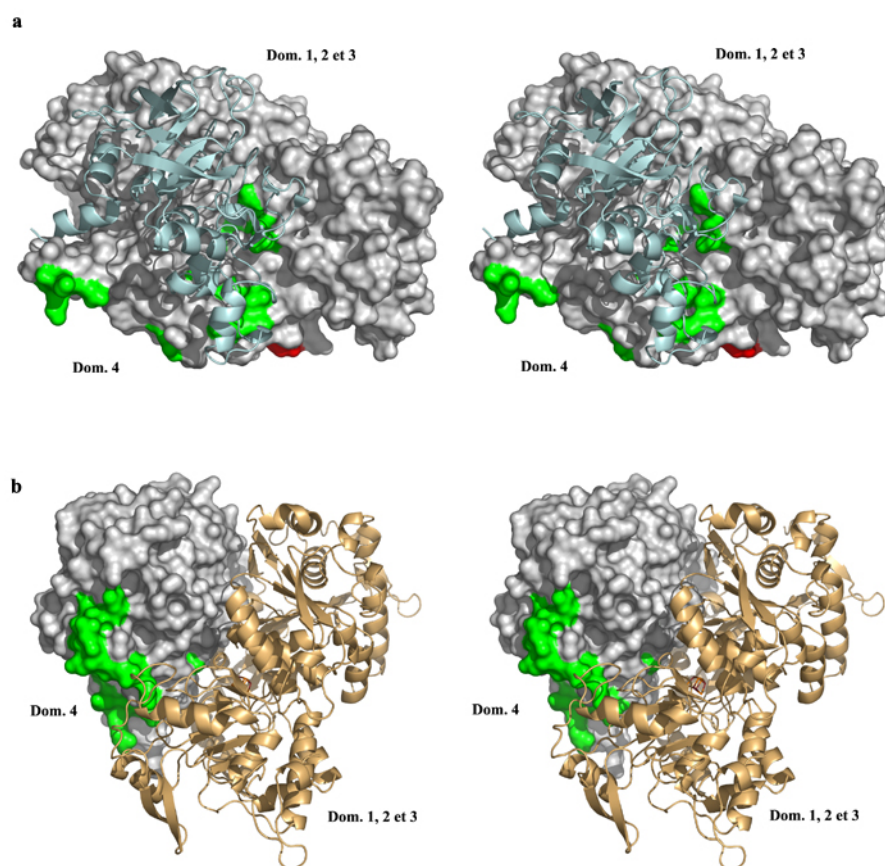


Figure IV.22 : Représentation des zones de contacts entre les domaines 1, 2, 3 et le domaine 4 sans le *linker*. Dans la figure a) le domaine 4 est représenté en bleu alors que dans le b) les domaines 1, 2 et 3 sont représentés en marron clair. Les régions de couleurs sont définies de la même manière qu'à la figure IV.21.

Tous les acides aminés étant décrits comme nécessaires à la fixation des motifs IRE sont, dans leur grande majorité, basiques. Ceci est facilement compréhensible du fait même de la fixation de l'ARN et de la présence dans sa structure de phosphates chargés négativement. En revanche certains acides aminés aromatiques sont impliqués dans les contacts entre les bases et la protéine (Jones et al., 2001). En effet, il a été démontré que l'effet hydrophobique et électrostatique des aromatiques permettaient des contacts attractifs d'une grande sélectivité ce qui est d'une importance capitale pour la reconnaissance et la spécificité des interactions ARN protéine (Waters, 2002). Ainsi des calculs d'accessibilité, réalisés avec le programme DSSP (Kabsch et Sander ; 1983), ont permis de mettre en évidence que, dans l'éventualité où le domaine 4 venait à effectuer un mouvement d'ouverture, plusieurs acides aminés

aromatiques des domaines 1, 2 et 3 se découvriraient au solvant. Ainsi, les phénylalanines 88, 399 et 538 et le tryptophane 219 sont tous plus exposés si le domaine 4 se déplace par rapport à la position qu'il occupe dans la forme aconitase d'IRP1. La phénylalanine 88 présente le plus gros écart d'accessibilité selon la présence ou non du domaine 4. Le tableau IV.2 présente une sélection des acides aminés aromatiques impliqués dans l'interface entre le domaine 4 et les domaines 1, 2 et 3 avec des distances inférieures à 5 Å.

AA hydrophobes	IRP1 (Å ²)	dom123 (Å ²)	Δ (Å ²)
Phe 88	10	105	95
Trp 219	5	17	12
Phe 399	102	110	8
Phe 538	46	62	16
Tyr 709	4	4	0
Trp 782	6	16	10
Tyr 873	17	52	35
Trp 882	52	65	13

Tableau IV.2 : Accessibilité des acides aminés aromatiques impliqués dans des contacts entre le domaine 4 et les domaines 1, 2 et 3 sur une distance inférieure à 5 Å. Sélection réalisée à partir du programme CONTACT (CCP4, 1994). Les calculs d'accessibilité ont été réalisés à partir du logiciel DSSP et représentent la surface du résidu exposé au solvant (Kabsch et Sander, 1983). Les acides aminés dont le numéro est inférieur à 660 correspondent aux domaines 1, 2 et 3 alors que au-delà ils font parti du domaine 4. Le dom123 correspond à la forme ouverte obtenue après séparation du domaine 4 et sans le linker.

Grâce à ce tableau il est possible de constater que les acides aminés aromatiques du domaine 4 ne sont pas dans leur ensemble très perturbés par une éventuelle ouverture de la gorge. Il n'y a que la tyrosine 873 qui se verrait un peu plus accessible que les autres mais pas de façon aussi drastique que pour la phénylalanine 88. L'ensemble de ces données vient s'ajouter aux contraintes déjà imposées par la structure d'IRP1 (Figure IV.23)

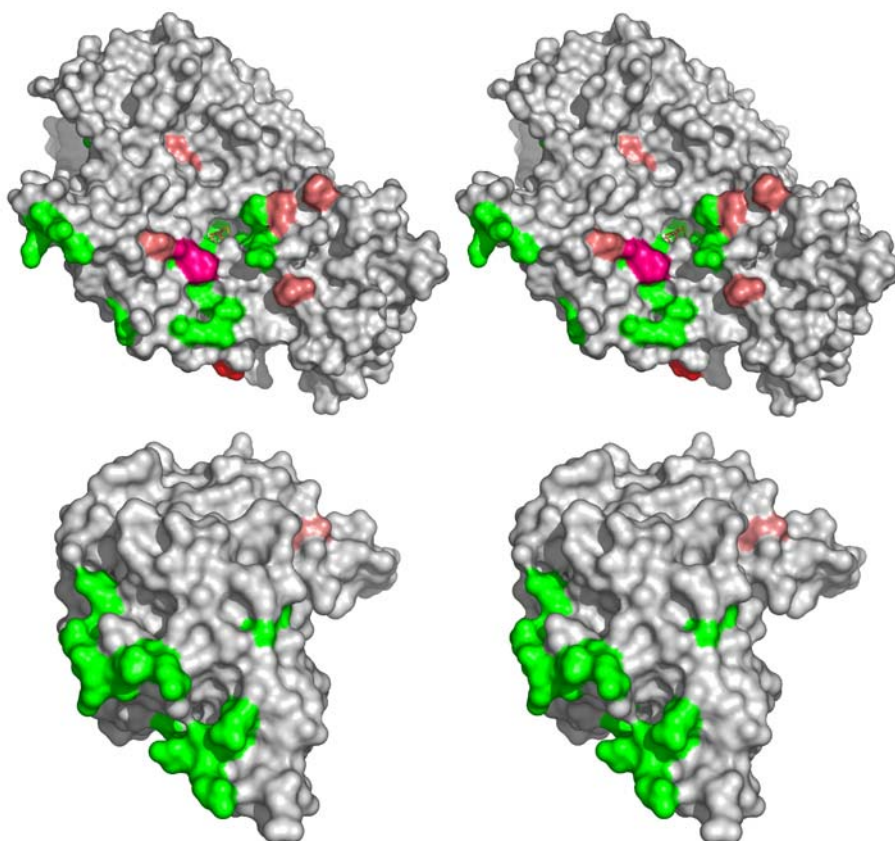


Figure IV.23 : Localisation des acides aminés supposés comme étant impliqués dans la fixation des motifs IRE. En haut, les domaines 1, 2 et 3 présentent en vert les acides aminés identifiés comme étant nécessaires ou protégés par la fixation de l'IRE, en rouge ceux qui sont surexposés, en saumon les acides aminés aromatiques dont l'accessibilité augmente lors d'un retrait du domaine 4 et en rose la phénylalanine 88. En bas les couleurs sont identiques au premier schéma mais transposées au domaine 4.

La superposition de la structure du motif IRE (Address et al, 1999) sur le domaine 4 montre clairement une similitude de forme (Figure IV.24). Dans la représentation de surface de la Figure IV.24, la couleur verte permet de localiser les acides aminés identifiés comme étant en interaction avec les IRE. En comparant spatialement cette localisation avec la figure IV.22, il est facile de constater que cette similitude est maximale pour la zone du domaine 4 qui est en interaction avec les domaines 1, 2 et 3 suggérant que l'ARN vient se loger à sa place sur la protéine. Une situation similaire a déjà été décrite dans la littérature avec comme exemple le plus marquant le facteur d'élongation EF-Tu qui voit un de ses domaines être complètement remplacé par un ARN de transfert avant de pouvoir être fonctionnel (Nissen et al., 1995). Il semble donc possible, dans un premier temps, d'essayer de positionner le motif IRE à la place du domaine 4 avant d'anticiper les mouvements que pourrait avoir ce même domaine après la fixation de l'ARN.

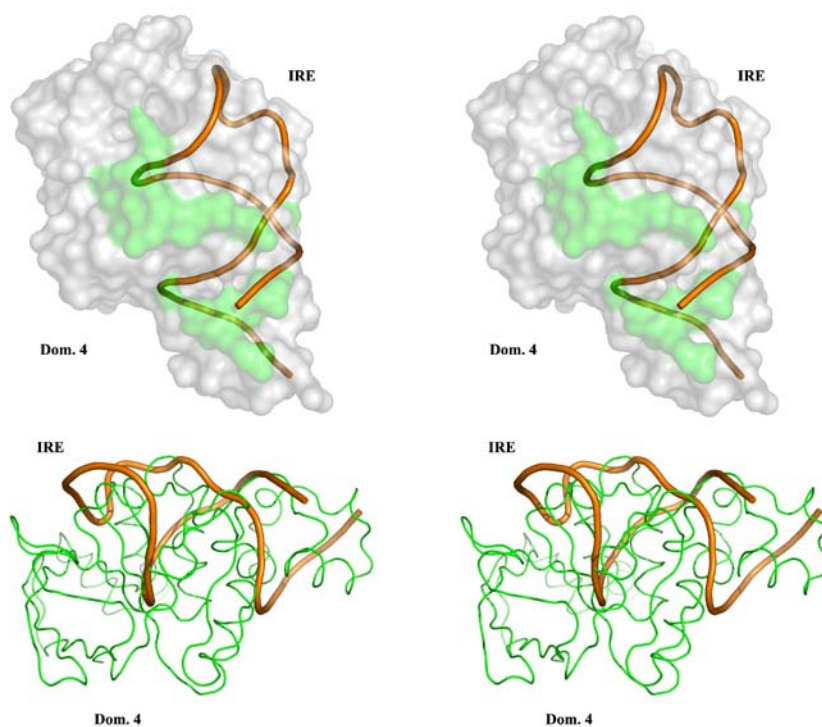


Figure IV.24 : Superposition du motif IRE et du domaine 4.

Il est rapidement apparu que deux orientations, tête bêche l'une par rapport à l'autre, étaient concevables pour le motif IRE à la surface des domaines 1, 2 et 3 (Figure IV.25).

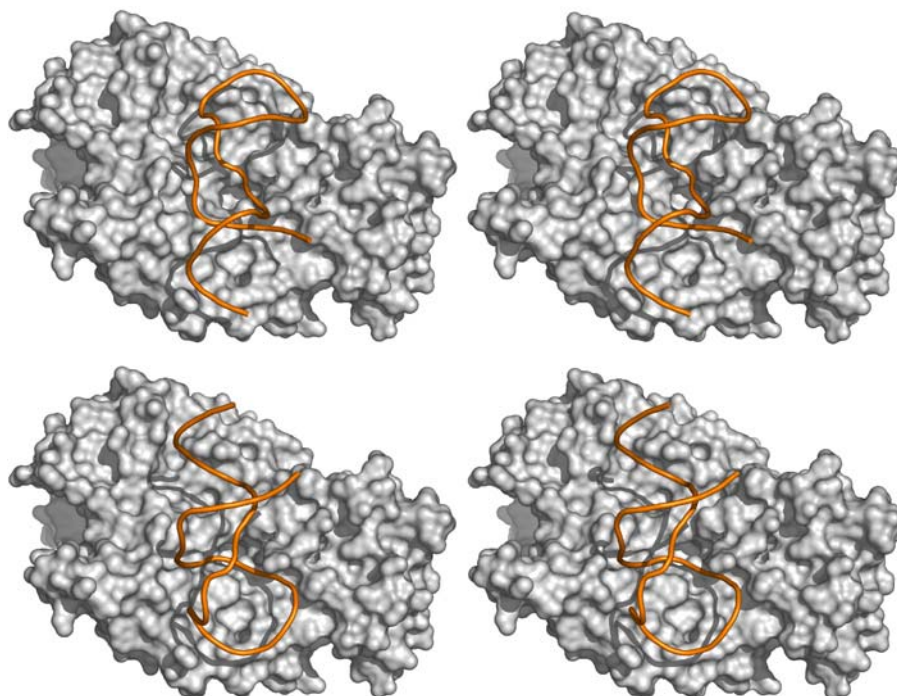


Figure IV.25 : Représentation des deux orientations possibles pour le positionnement de l'IRE à la surface des domaines 1, 2 et 3. L'orientation du motif IRE sera qualifiée arbitrairement de « tête haute » lorsque la boucle de l'ARN sera positionnée comme sur le schéma du haut et de « tête basse » lorsque la boucle sera inversée comme illustré sur le schéma du bas.

En nous appuyant sur le travail réalisé par Paraskeva et al. (1996) notre choix s'est rapidement orienté sur la forme « tête haute » comme présenté sur la Figure IV.25. Afin de parfaire la modélisation, il fallait encore vérifier que tous les acides aminés identifiés comme étant en interaction avec l'IRE étaient réellement à proximité de l'ARN. Il va de soi que cette modélisation ne peut pas retranscrire la réorganisation de certaines structures secondaires. Il n'est donc pas envisageable de pouvoir expliquer tous les contacts par cette seule approche mais cela peut permettre d'élucider certaines hypothèses encore non avérées. Comme le présente la Figure IV.26 le motif « tête haute » semble pouvoir interagir avec bon nombre d'acides aminés décrits comme faisant des liaisons ou des contacts avec l'IRE.

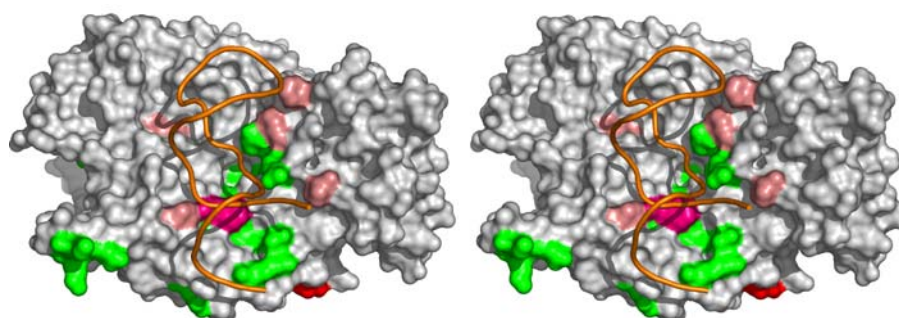


Figure IV.26 : Positionnement du motif IRE forme « tête haute » sur les domaines 1, 2 et 3. En vert les acides aminés identifiés comme étant nécessaires ou protégés par la fixation de l'IRE, en rouge ceux qui sont surexposés, en saumon les acides aminés aromatiques dont l'accessibilité augmente lors d'un retrait du domaine 4 et en rose la phénylalanine 88.

Il restait alors à placer le domaine 4 de telle sorte que les acides aminés, là encore, nécessaires à la fixation de l'IRE puissent être idéalement positionnés. Il fallait ensuite s'assurer que les nucléotides identifiés comme étant protégés le soient vraiment avant de valider notre modèle (Figure IV.27 a). Même si le modèle semble être de bonne qualité une rapide analyse montre que deux éléments jouent en sa défaveur. Pour commencer les nucléotides C7, A14, G15 et U16 sont bien protégés mais pas par les zones initialement attendues et décrites par Kaldy et al. (1999). En effet, il n'y a ni plus ni moins qu'une inversion des interactions prédites par les études de mutagenèse car la région Asn 685 – Asn 689 a été décrite comme rentrant en contact avec la boucle de l'IRE alors que la région Arg 732 – Gln 737 serait plutôt en interaction avec la base non appariée, ce qui n'est pas le cas dans notre modélisation (Figure IV.27 b).

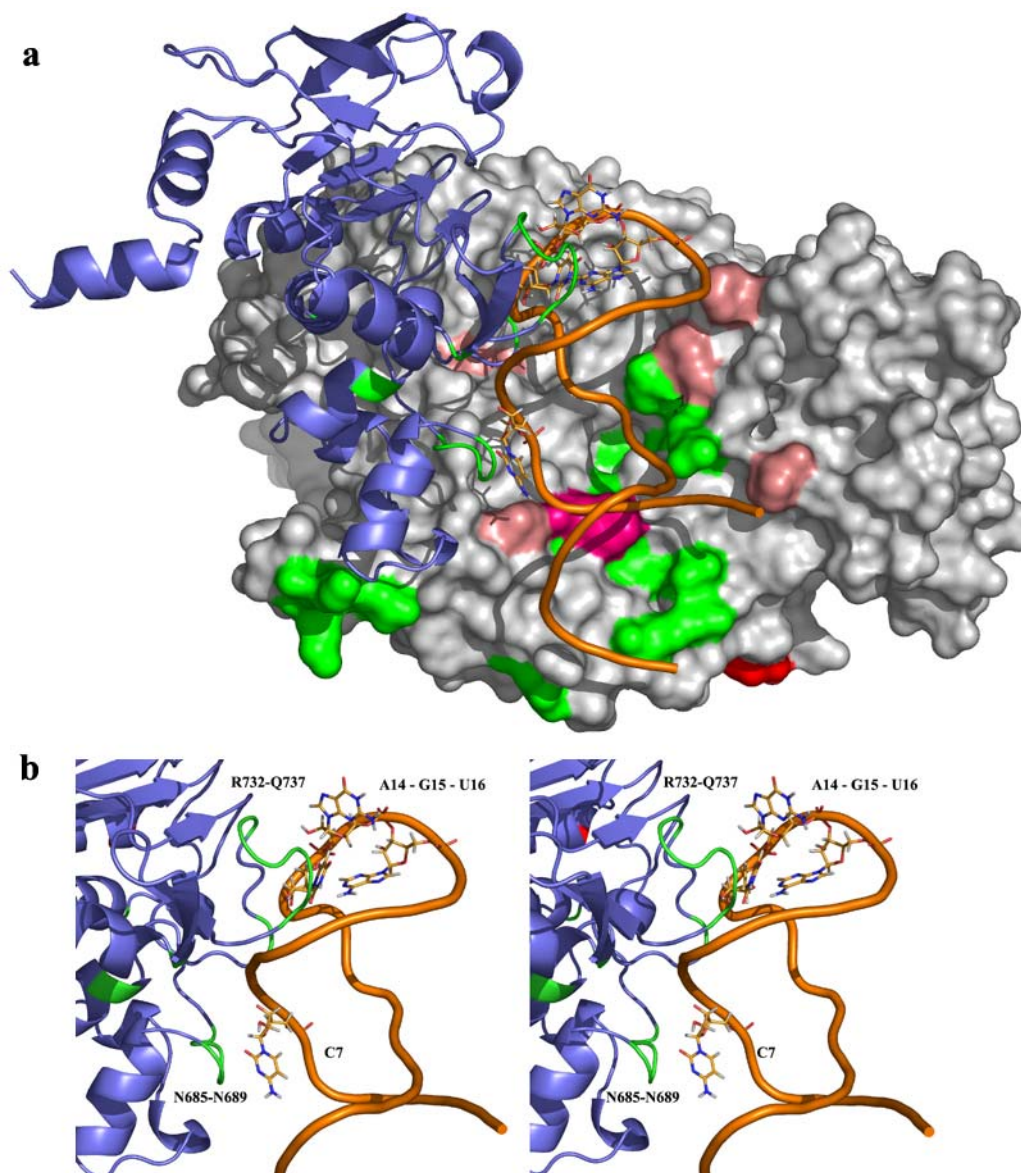


Figure IV.27 : Modèle final avec un IRE en position tête haute. En bleu est représenté le domaine 4. a) Vue générale de la modélisation et en b) zoom sur les contacts entre le domaine 4 et l'IRE.

Un autre point qui va à l'encontre de ce modèle réside dans le fait que, contrairement à l'aconitase mitochondriale, les domaines 1 et 4 forment, de façon très localisée, un patch hydrophobe en mettant en commun certains de leurs acides aminés. Ces résidus, au nombre de 13, sont localisés sur les hélices $\alpha 4$ et $\alpha 7$ du domaine 1 et sur l'hélice $\alpha 31$ du domaine 4, à l'exception de l'un d'entre eux qui fait suite à l'hélice $\alpha 29$ (Figure IV.28). L'ensemble de ces résidus, Met 100, Leu 107, Ile 224, Ala 228, Leu 231, Met 811, Ile 879, Leu 880, Met 883, Ile 884, Met 887 et Ala 888, peut très bien former une charnière autour de laquelle le domaine 4 pourrait bouger.



Figure IV.28 : Localisation du patch hydrophobe constitué par le domaine 1 et le domaine 4. En haut une vue générale de la structure et en bas un zoom sur la région hydrophobe formée par les hélice $\alpha 4$, $\alpha 7$ et $\alpha 31$

Dans le modèle proposé le patch hydrophobe est clairement exposé au solvant ce qui n'est pas raisonnable (Figure IV.29). Paraskeva et al. (1996) ne pouvait pas prévoir ce genre de désagrément car l'aconitase mitochondriale n'a seulement que deux acides aminés de conservés sur la totalité de ceux nécessaires à la mise en place de ce patch.

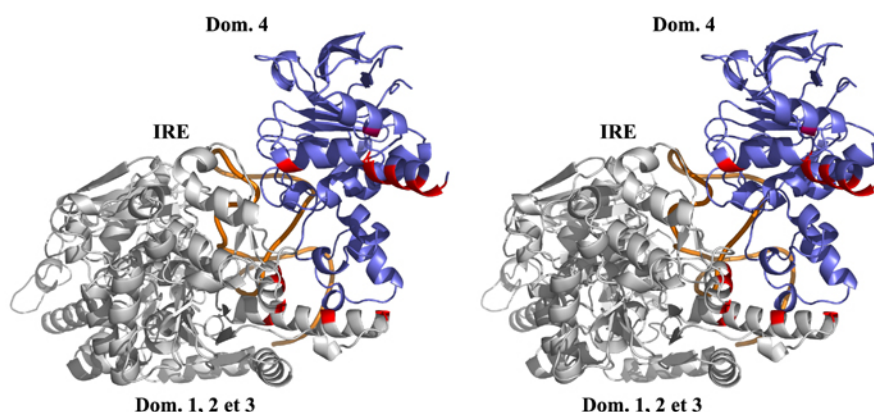


Figure IV.29 : Localisation de la zone hydrophobe constituée par les domaines 1 et 4 après la modélisation de la fixation du motif IRE.

Une nouvelle modélisation a donc été envisagée mais cette fois avec le motif IRE en position tête basse. La démarche a été identique à la précédente à savoir que nous avons commencé par prendre en considération la complémentarité des formes du motif IRE et des domaines 1, 2 et 3 ensuite la position des acides aminés protégés lors de la fixation de l'ARN nous a permis d'affiner le placement de l'IRE sur la protéine IRP1 (Figure IV.30).

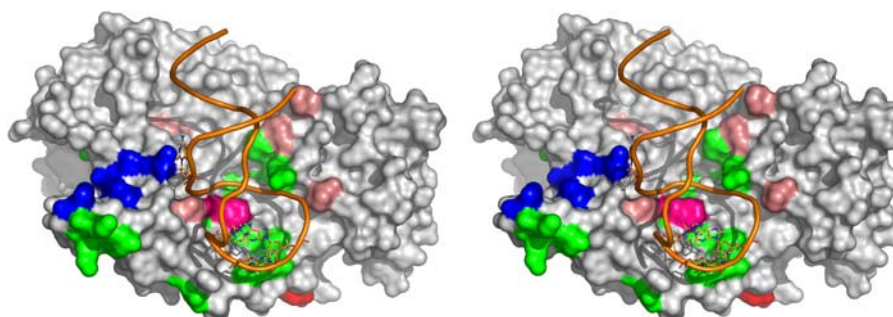


Figure IV.30 : Positionnement du motif IRE sous une forme tête basse. La zone de la charnière hydrophobe a été colorée en bleu, le reste des couleurs représentant toujours les mêmes choses que dans les figures précédentes.

Dans un deuxième temps, nous avons déplacé le domaine 4 de telle façon que les régions impliquées dans des contacts avec l'IRE puissent jouer pleinement leur rôle. Pour cette étape, nous avons tenu compte de la contrainte imposée par la charnière hydrophobe. Afin de satisfaire cette condition nous avons dû segmenter le domaine 4 en trois parties car les deux hélices α_{30} et α_{31} , cette dernière étant impliquée dans la charnière hydrophobe, ne pouvaient pas être modélisées en corps rigide avec le reste du domaine. Nous avons donc décidé de fixer la position de l'hélice α_{31} . Le domaine 4 a été placé par rapport à l'IRE et l'hélice α_{30} a été modifiée afin de pouvoir faire le lien entre la charnière et le reste de la protéine. Pour limiter les risques d'erreur, par rapport à la nouvelle position qu'occuperait l'hélice α_{30} , la tyrosine 873 nous a servi de point de repère et nous avons tenté de minimiser son déplacement lors de notre modélisation (Figure IV.31).

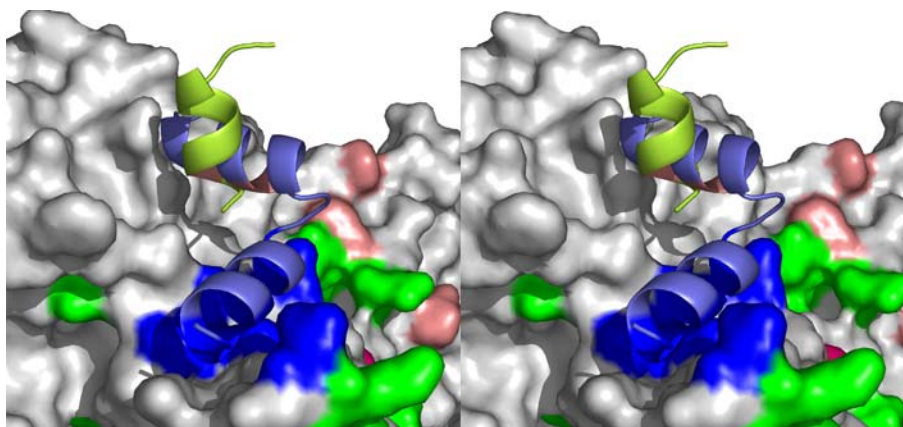


Figure IV.31 : Superposition de l'hélice $\alpha 30$ avant et après modélisation. Les hélices $\alpha 30$ et $\alpha 31$ dans leur position d'origine sont colorées en mauve. Sur l'hélice $\alpha 31$ la couleur bleue représente la complémentarité de la charnière hydrophobe alors que sur l'hélice $\alpha 30$ la couleur saumon permet de localiser la position de la tyrosine 873.

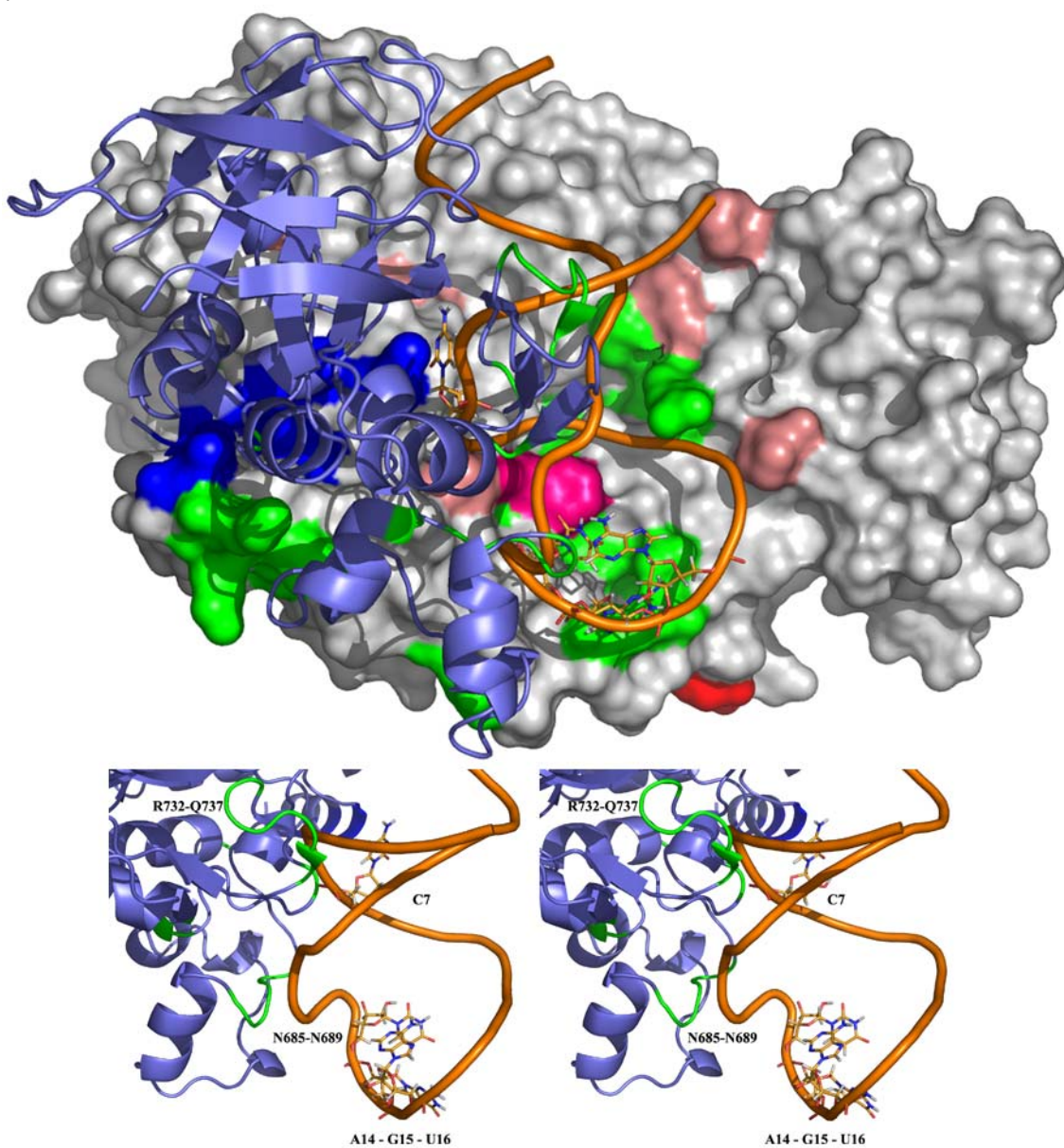


Figure IV.32 : Modélisation de la fixation de l'IRE sur IRP1. En bas zoom sur les interactions ARN-domaine 4.

D'une manière générale cette modélisation répond plus précisément aux contraintes que nous nous étions fixées (Figure IV.32). En effet, le positionnement de l'ARN sur la protéine semble être en accord avec l'identification des acides aminés décrits comme étant protégés dans la littérature. De plus, les régions interagissant avec la boucle et le « *bulge* » sont idéalement placées pour pouvoir effectuer des contacts avec l'ARN. Il est clair néanmoins que la flexibilité de ces régions ainsi que celle de l'IRE ne peuvent que permettre un meilleur contact que celui qui est représenté au bas de la figure IV.32. En revanche, ce modèle ne permet pas d'expliquer pourquoi l'aspartate 137 et la lysine 141 se retrouvent surexposés aux protéases après la fixation de l'IRE sur IRP1. Il est aussi intéressant de constater que les acides aminés aromatiques, et donc hydrophobes, des domaines 1, 2 et 3 sont tous protégés soit par la fixation de l'IRE soit par le déplacement du domaine 4. Ce même déplacement permet aussi d'expliquer que certains acides aminés, qui n'étaient pas protégés lors de la première modélisation, le sont maintenant à savoir les arginines 101 et 102, les lysines 105 et 106 et le glutamate 155 (voir Figure IV.21 a). Afin de conclure sur cette modélisation il nous restait à expliquer pourquoi la lysine 79, l'aspartate 121 et l'arginine 187 avaient été décrites comme étant préservées d'une attaque protéolytique lors de la fixation d'un IRE. Il est à noter que ces acides aminés ne sont pas du tout présents sur la surface rentrant en contact avec l'ARN mais au contraire qu'ils sont localisés sur la face opposée à cette interaction comme la figure IV.21 b le montre.

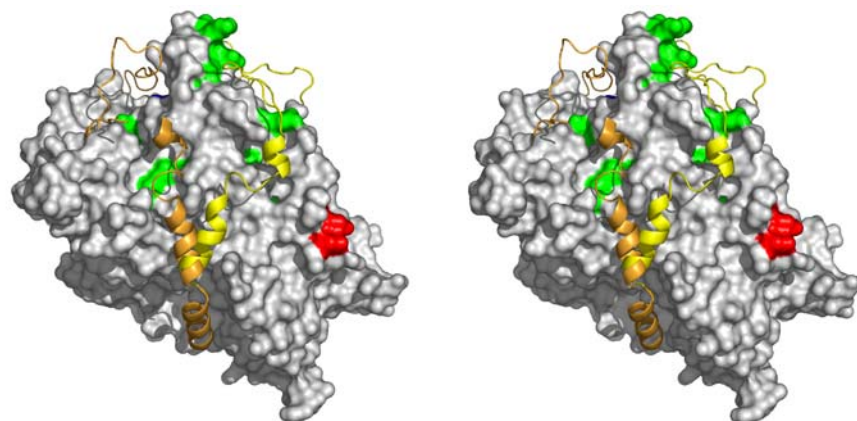


Figure IV.33 : Modélisation de la nouvelle position du *linker*. Le jaune représente la position qu'adopte le *linker* dans la protéine IRP1 et l'orange symbolise le linker modélisé lors de la fixation d'un IRE sur IRP1.

En se focalisant un peu plus sur cette région, il est possible d'imaginer que le *linker* subit, au même titre que le reste de la protéine, un remaniement dans son positionnement par rapport aux domaines 1, 2 et 3. Basé sur une complémentarité des formes, sur de possibles interactions entre les acides aminés chargés du *linker* et les trois acides aminés protégés et enfin, en tenant compte de la longueur nécessaire pour rattacher le domaine 4 aux domaines 1,

2 et 3, il est possible de spéculer sur ses mouvements et sur sa nouvelle position (Figure IV.33). La structure du *linker* est tellement flexible que nous pouvons accorder qu'un crédit relatif à cette modélisation mais cela semble pouvoir expliquer pourquoi certains acides aminés ont été décrits comme étant importants pour la fixation de l'ARN alors qu'ils ne sont pas directement en contact avec l'IRE.

Basé sur l'ensemble de ces considérations, il est possible de penser que notre modélisation est actuellement la seule à être en mesure d'expliquer tout ce que la littérature a pu recenser comme étant nécessaire aux interactions entre l'ARN et la protéine. Toutes les modélisations précédemment réalisées avec l'aconitase mitochondriale de bœuf ne pouvaient que suggérer des idées sur le type de remaniements structuraux que subissait IRP1 mais sans tenir compte des différences qui existent entre IRP1 et mAcn. Notre modèle semble être en mesure de servir de point d'appui pour de futures études de mutagenèse dirigée.

3 IRP1 versus IRP2

IRP1 et IRP2 sont intimement liées l'une à l'autre. Toutes les deux sont impliquées dans le métabolisme du fer mais à des degrés différents de complexité et de fonctionnement. IRP2 semble être « la » protéine régulatrice du fer responsable du contrôle de son homéostasie cellulaire alors qu'IRP1 jouerait un rôle plus spécialisé dans le cadre de la réponse au stress oxydant via une régulation de la disponibilité du fer dans la cellule. IRP1 et IRP2 partagent près de 61% d'identité et environ 80% d'homologie de séquence (Figure IV.34). Sur l'ensemble des 21 acides aminés impliqués dans la fonction aconitase d'IRP1 seulement 18 sont conservés par la protéine IRP2. Les trois cystéines impliquées dans la coordination de l'agrégat fer soufre sont présentes de même que les deux asparagines, Asn 298 et 535, qui forment des liaisons avec les soufres voisins mais en revanche l'isoleucine 507 qui forme un contact hydrophobe avec l'agrégat est remplacée par une valine. L'encombrement stérique de cet acide aminé est moins important que celui d'une isoleucine mais il conserve néanmoins toutes les propriétés hydrophobes nécessaires à un contact avec l'agrégat. Ainsi de prime abord, IRP2 semble présenter toutes les conditions requises pour pouvoir incorporer un agrégat fer soufre et ainsi acquérir la fonction aconitase mais ce n'est pourtant pas le cas. Si l'analyse est poursuivie un peu plus en avant il apparaît que la sérine 778 d'IRP1 est remplacée par une asparagine dans IRP2. Cette sérine est la base du mécanisme catalytique de l'isomérisation du citrate en isocitrate (voir § IV.1.3.2).

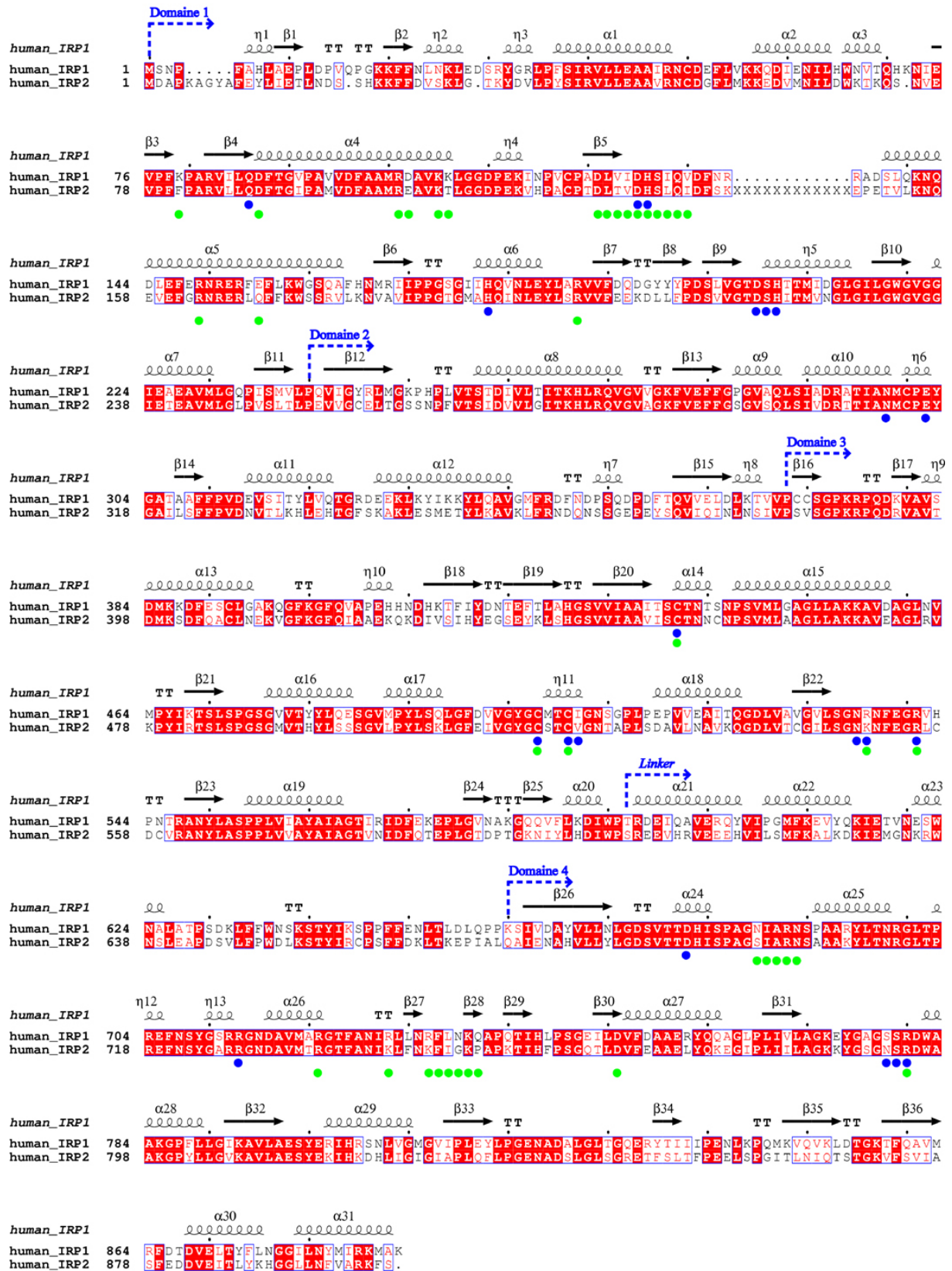


Figure IV.34 : Alignement de séquences entre la protéine IRP1 humaine et IRP2. Les XXX correspondent à la zone où les 73 acides aminés supplémentaires d'IRP2 s'insèrent. Les points bleus symbolisent les acides aminés nécessaires à l'activité aconitase d'IRP1 alors que les points verts représentent les acides aminés identifiés comme étant impliqués dans la fixation des motifs IRE chez IRP1.

L'absence de ce résidu prive IRP2 de toute activité catalytique ; tout du moins si la structure tridimensionnelle est conservée entre IRP1 et IRP2 et si elle avait été capable de fixer un agrégat fer soufre. Il est aussi intéressant de noter que la sérine 441 présente dans IRP1 est mutée en cystéine dans IRP2 (Figure IV.35). Il est donc tout à fait possible d'envisager que cette cystéine puisse venir former un pont disulfure avec la cystéine 437, si elle adopte une autre conformation, empêchant ainsi toutes possibilités de coordination d'un agrégat fer soufre. Il en va de même pour l'isoleucine 176 qui est mutée en méthionine dans IRP2, ce qui pourrait avoir comme conséquence d'empêcher l'insertion d'un agrégat fer soufre par obstruction (Figure IV.35).

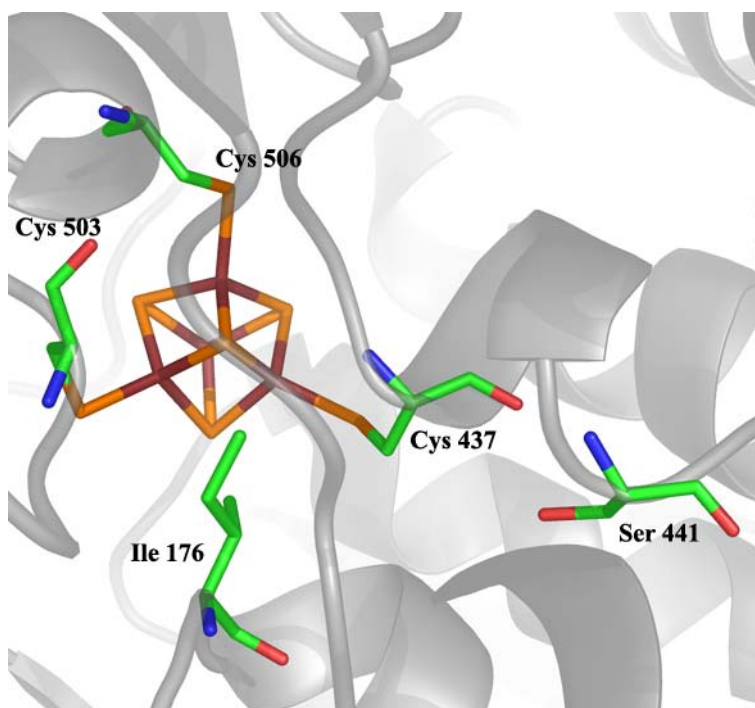


Figure IV.35 : Localisation de la sérine 441 et de l'isoleucine 176 à proximité de l'agrégat fer soufre.

La non activité aconitase d'IRP2 pourrait aussi dépendre d'un problème d'accessibilité du substrat au cœur catalytique de la protéine. En effet, le citrate au pH physiologique est chargé négativement (voir § IV.1.3.3) ce qui lui impose d'avoir des interactions avec des résidus basiques pour pouvoir être conduit jusqu'à l'agrégat fer soufre. Or, cette seule hypothèse n'est pas valide car tous les acides aminés tapissant les tunnels d'accessibilités dans IRP1 sont présents dans IRP2. Il est aussi important de noter que IRP2 présente une insertion de 73 acides aminés par rapport à IRP1. Cette insertion a lieu dans le domaine 1 par rapport à un alignement à IRP1 et à la surface de la protéine comme schématisé sur la figure IV.36. C'est par ce domaine que IRP2 fixerait un hème ce qui aurait pour conséquence de conduire la protéine vers sa dégradation par le protéasome.

L'analyse des résidus, ayant été décrits comme impliqués dans la fixation des motifs IRE, montre que sur les 39 résidus d'IRP1 seulement 25 sont conservés chez IRP2. Une différence notable se fait sur les acides aminés 732 à 735 qui est la région décrite comme interagissant avec la cytosine non appariée des IRE. Sur ces six résidus seulement deux d'entre eux sont conservés, ce qui laisse à penser que IRP2 n'interagit peut-être pas de la même manière que IRP1 avec les motifs IRE, surtout si l'ensemble des différences entre IRP1 et IRP2 est pris en considération.

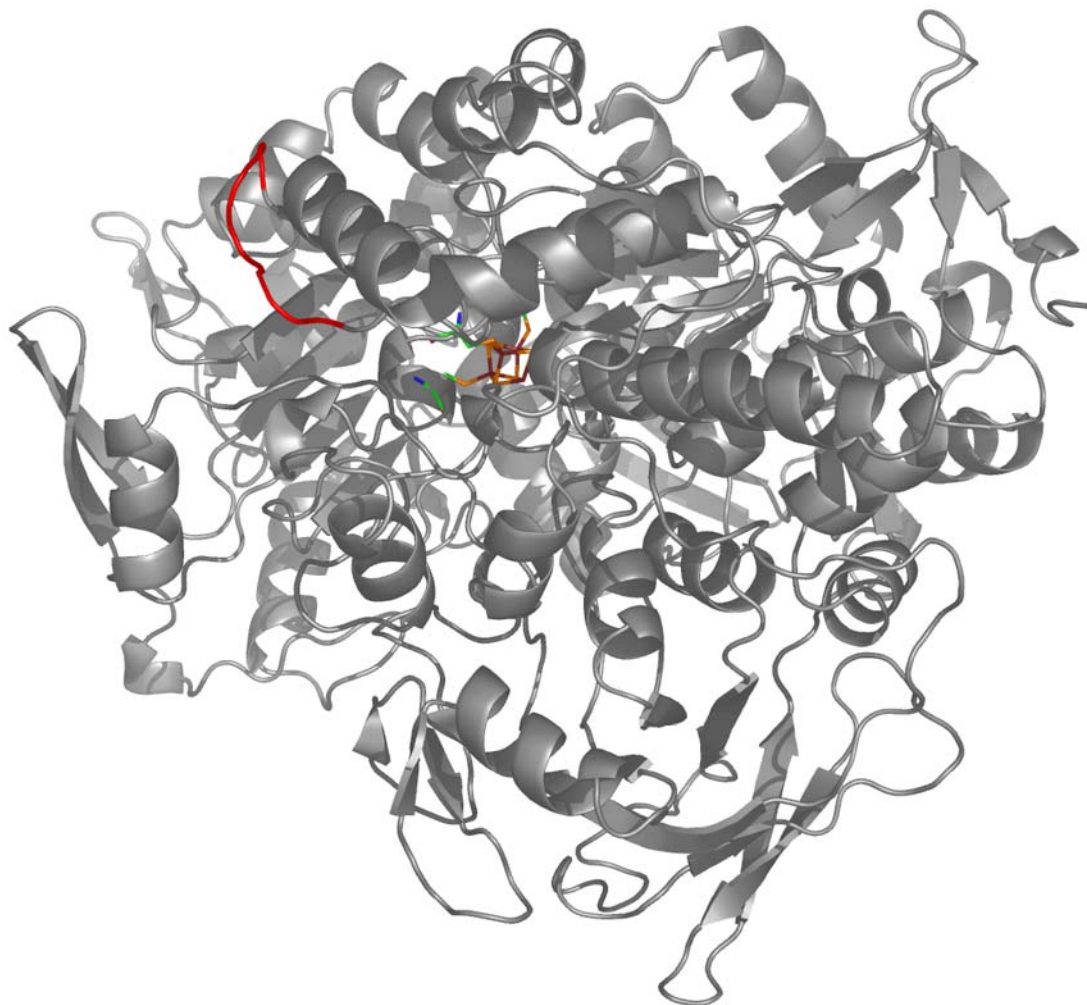


Figure IV.36 : Localisation de la zone d'insertion des 73 résidus supplémentaires de IRP2 sur la structure d'IRP1. Le rouge symbolise la zone où l'insertion doit avoir lieu.

D'une manière générale, il est possible de s'attendre à ce que la structure d'IRP2 soit relativement similaire à celle d'IRP1 néanmoins il est intéressant de la résoudre ne serait ce que pour les 73 résidus de l'insertion et sur le mode de fixation de l'hème. A l'heure actuelle, le mode de dégradation d'IRP2 semble être associé à la présence de cet insert mais les travaux, même les plus récents, ne tranchent pas définitivement sur l'importance ou non de sa présence. Il paraît donc clair que la résolution de la structure tridimensionnelle d'IRP2

pourrait apporter une aide non négligeable à la compréhension du mode de fonctionnement d'IRP2.

Conclusion et perspectives



Chapitre V : Conclusion et perspectives

L'homéostasie du fer est essentielle à tout les métazoaires et tout particulièrement à l'homme. Une dérégulation de ses mécanismes de contrôle aboutit à des conséquences désastreuses sur l'individu, pouvant entraîner la mort *in fine*. IRP1 et IRP2 sont deux protéines au cœur de cette machinerie de contrôle que ce soit pour une simple cellule ou, d'une manière plus générale, pour l'ensemble de l'organisme. IRP1 semble être plutôt spécialisée dans le contrôle du stress oxydant. Il est de plus en plus évident, qu'au travers de la modulation du niveau de fer libre dans la cellule, IRP1 peut faire varier le stress imposé à une cellule ou à un tissu. Ceci est parfaitement illustré par les phénomènes inflammatoires. En effet, IRP1 est étroitement lié à l'activation des neutrophiles et autres macrophages lors de la défense de l'organisme face à un pathogène au travers de la production d'espèces très réactives dérivées de l'oxygène. Ainsi, sous l'action d'agents oxydants tels que le peroxyde d'hydrogène, IRP1 est capable, après phosphorylation (voir § I.2.1.1.1), de perdre son agrégat fer soufre et d'acquérir son activité de fixation à l'ARN. Cette augmentation d'activité n'a jamais été détectée pour IRP2 ce qui tendrait à prouver la spécificité d'IRP1 pour le stress oxydant. Cette activation entraîne une stabilisation des ARNm du TfR et une diminution du temps de demi vie des ARNm de la Ft. La conséquence est une augmentation de l'incorporation de fer au sein de la cellule et donc d'une production accrue d'hydroxyles et d'autres espèces hautement réactives suivant le principe de la réaction de Fenton (voir § I.1). Cette cascade oxydative initiée par les leucocytes est un moyen efficace de lutter contre les pathogènes. Il semble donc de plus en plus probable qu'au détriment de son « simple » rôle de régulateur de l'homéostasie du fer, IRP1 puisse être un élément central de la stabilité des niveaux redox au niveau de certaines cellules pour ne pas dire de toutes.

Ce travail a permis de mettre en évidence un parfait maintien de la fonction aconitase de la protéine IRP1 avec la conservation de tous les acides aminés nécessaires à la réaction. Au-delà de cette seule conservation, la topologie du cœur catalytique a parfaitement été préservée, ajoutant plus de crédit à un mode de fonctionnement identique à celui de l'aconitase mitochondriale. Là encore, l'activité isomérase d'IRP1 semble être en adéquation avec les variations redox de la cellule et ce au travers du système « glutathion » via le contrôle de la concentration en NADPH cytoplasmique.

La structure d'IRP1 a aussi permis d'émettre quelques hypothèses quant aux effets de la phosphorylation des différentes cibles localisées sur la protéine (voir § IV.1.3.5). Dans les

deux cas, les sérines sont à proximité du site catalytique d'IRP1 et leur phosphorylation semble pouvoir avoir des effets directs sur l'activité de la protéine ce qui est en accord avec la littérature. De plus la modélisation de la fixation de l'IRE sur IRP1 a permis de dégager quelques idées sur ce que pourrait être la structure du complexe. En s'appuyant sur ce modèle, certaines conclusions ont pu être apportées concernant les régions de la protéine impliquées dans l'interaction avec l'ARN tout en tenant compte des précédents résultats accumulés au fil des années dans la littérature scientifique couvrant le sujet. C'est ainsi qu'il apparaît que certains acides aminés, décrits comme nécessaire à la fixation des motifs IRE, se trouvent en réalité localisés sur la face opposée à la surface rentrant en contact avec l'ARN. Il paraît maintenant clair que leur interaction est indirecte sur les motifs IRE mais sûrement nécessaire quant à la stabilité de la structure du complexe.

D'une manière générale, il reste encore quelques points à éclaircir concernant IRP1. Tout d'abord la forme à plus haute résolution, en P2₁, présente une légère différence entre les deux molécules de l'unité asymétrique. En effet, le domaine 4 semble avoir adopté une conformation légèrement plus ouverte sur une des deux molécules. Ceci semble être en corrélation avec le fait que la molécule fermée possède, dans son site actif, une molécule de substrat dont l'occupation ne nous permet pas de trancher de façon définitive entre le *cis*-aconitate et l'isocitrate. De nouvelles expériences devraient permettre d'obtenir des cristaux a même de répondre à ce genre de question. La présence de l'intermédiaire réactionnel, le *cis*-aconitate, serait une première car aucune structure n'a actuellement été réalisée en sa présence. Cela pourrait être l'occasion de trancher sur le mode de fixation de cet intermédiaire sur l'agrégat fer soufre ainsi que de confirmer l'existence du « *flip* » nécessaire à la réaction.

Des essais de cristallisation du complexe sont actuellement en cours au sein du laboratoire. Toujours dans le cadre d'une meilleure compréhension des mécanismes de régulation de l'homéostasie du fer chez l'homme et afin d'élargir notre champs d'activité, IRP2 semble être une suite logique à la continuité de ce projet. La structure de cette protéine permettrait une meilleure compréhension des différences qui existent entre IRP1 et IRP2. Cela répondrait aussi, de façon définitive, au pourquoi de l'absence d'activité aconitase pour cette protéine. De même la structure de l'insert de 73 acides aminés pourrait amener une certaine lumière quant au mode de dégradation d'IRP2 qui est actuellement encore mal défini.

Bibliographie



Bibliographie

A

Abboud, S. and Haile, D.J. (2000) A novel mammalian iron-regulated protein involved in intracellular iron metabolism. *J Biol Chem*, 275, 19906-19912.

Abrahams, J.P. and Leslie, A.G.W. (1996) Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Cryst.*, D52, 30-42.

Adamec, J., Rusnak, F., Owen, W.G., Naylor, S., Benson, L.M., Gacy, A.M. and Isaya, G. (2000) Iron-dependent self-assembly of recombinant yeast frataxin: implications for Friedreich ataxia. *Am J Hum Genet*, 67, 549-562.

Address, K.J., Basilion, J.P., Klausner, R.D., Rouault, T.A. and Pardi, A. (1997) Structure and dynamics of the iron responsive element RNA: implications for binding of the RNA by iron regulatory binding proteins. *J Mol Biol*, 274, 72-83.

Allerson, C.R., Cazzola, M. and Rouault, T.A. (1999) Clinical severity and thermodynamic effects of iron-responsive element mutations in hereditary hyperferritinemia-cataract syndrome. *J Biol Chem*, 274, 26439-26447.

Allikmets, R., Raskind, W.H., Hutchinson, A., Schueck, N.D., Dean, M. and Koeller, D.M. (1999) Mutation of a putative mitochondrial iron transporter gene (ABC7) in X-linked sideroblastic anemia and ataxia (XLSA/A). *Hum Mol Genet*, 8, 743-749.

Aloni, R., Peleg, D. and Meyuhas, O. (1992) Selective translational control and nonspecific posttranscriptional regulation of ribosomal protein gene expression during development and regeneration of rat liver. *Mol Cell Biol*, 12, 2203-2212.

Audic, S., Lopez, F., Claverie, J.M., Poirot, O. and Abergel, C. (1997) SAmBA: an interactive software for optimizing the design of biological macromolecules crystallization experiments. *Proteins*, 29, 252-257.

B

Balla, G., Jacob, H.S., Balla, J., Rosenberg, M., Nath, K., Apple, F., Eaton, J.W. and Vercellotti, G.M. (1992) Ferritin: a cytoprotective antioxidant strategem of endothelium. *J Biol Chem*, 267, 18148-18153.

Bartzokis, G. (2004) Age-related myelin breakdown: a developmental model of cognitive decline and Alzheimer's disease. *Neurobiol Aging*, 25, 5-18; author reply 49-62.

- Basilion, J.P., Rouault, T.A., Massinople, C.M., Klausner, R.D. and Burgess, W.H. (1994) The iron-responsive element-binding protein: localization of the RNA-binding site to the aconitase active-site cleft. *Proc Natl Acad Sci U S A*, 91, 574-578.
- Beinert, H., Holm, R.H. and Munck, E. (1997) Iron-sulfur clusters: nature's modular, multipurpose structures. *Science*, 277, 653-659.
- Beinert, H., Kennedy, M.C. and Stout, C.D. (1996) Aconitase as Ironminus signSulfur Protein, Enzyme, and Iron-Regulatory Protein. *Chem Rev*, 96, 2335-2374.
- Beinert, H. and Kiley, P. (1996) Redox control of gene expression involving iron-sulfur proteins. Change of oxidation-state or assembly/disassembly of Fe-S clusters? *FEBS Lett*, 382, 218-219; discussion 220-211.
- Blundell, T.L. and Johnson, L.N. (1976) *Protein Crystallography*. Academic Press Inc., New-York.
- Bodovitz, S., Falduto, M.T., Frail, D.E. and Klein, W.L. (1995) Iron levels modulate alpha-secretase cleavage of amyloid precursor protein. *J Neurochem*, 64, 307-315.
- Bouton, C., Hirling, H. and Drapier, J.C. (1997) Redox modulation of iron regulatory proteins by peroxynitrite. *J Biol Chem*, 272, 19969-19975.
- Bouton, C., Oliveira, L. and Drapier, J.C. (1998) Converse modulation of IRP1 and IRP2 by immunological stimuli in murine RAW 264.7 macrophages. *J Biol Chem*, 273, 9403-9408.
- Bouton, C., Raveau, M. and Drapier, J.C. (1996) Modulation of iron regulatory protein functions. Further insights into the role of nitrogen- and oxygen-derived reactive species. *J Biol Chem*, 271, 2300-2306.
- Braak, H., Braak, E. and Bohl, J. (1993) Staging of Alzheimer-related cortical destruction. *Eur Neurol*, 33, 403-408.
- Brazzolotto, X., Gaillard, J., Pantopoulos, K., Hentze, M.W. and Moulis, J.M. (1999) Human cytoplasmic aconitase (Iron regulatory protein 1) is converted into its [3Fe-4S] form by hydrogen peroxide in vitro but is not activated for iron-responsive element binding. *J Biol Chem*, 274, 21625-21630.
- Brazzolotto, X., Timmins, P., Dupont, Y. and Moulis, J.M. (2002) Structural changes associated with switching activities of human iron regulatory protein 1. *J Biol Chem*, 277, 11995-12000.
- Brittenham, G.M. (1994) The red cell cycle. In *iron metabolism in health and disease*. J.H. Brock et al., eds. (London: W.B. Saunders Company LTD.), 31-62.
- Brown, N.M., Anderson, S.A., Steffen, D.W., Carpenter, T.B., Kennedy, M.C., Walden, W.E. and Eisenstein, R.S. (1998) Novel role of phosphorylation in Fe-S cluster stability revealed by phosphomimetic mutations at Ser-138 of iron regulatory protein 1. *Proc Natl Acad Sci U S A*, 95, 15235-15240.
- Brun, A. and Englund, E. (1986) A white matter disorder in dementia of the Alzheimer type: a pathoanatomical study. *Ann Neurol*, 19, 253-262.

Brünger, A.T. (1992a) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, 355, 475.

Brünger, A.T. (1992b) X-PLOR version 3.1. A system for X-ray crystallography and NMR. Yale Univ. Press, New Haven, CT.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallogr D Biol Crystallogr*, 54, 905-921.

Burdon, R.H. (1996) Control of cell proliferation by reactive oxygen species. *Biochem Soc Trans*, 24, 1028-1032.

Butt, J., Kim, H.Y., Basilion, J.P., Cohen, S., Iwai, K., Philpott, C.C., Altschul, S., Klausner, R.D. and Rouault, T.A. (1996) Differences in the RNA binding sites of iron regulatory proteins and potential target diversity. *Proc Natl Acad Sci U S A*, 93, 4345-4349.

C

Cairo, G., Castrusini, E., Minotti, G. and Bernelli-Zazzera, A. (1996) Superoxide and hydrogen peroxide-dependent inhibition of iron regulatory protein activity: a protective stratagem against oxidative injury. *Faseb J*, 10, 1326-1335.

Cairo, G. and Pietrangelo, A. (1994) Transferrin receptor gene expression during rat liver regeneration. Evidence for post-transcriptional regulation by iron regulatory factorB, a second iron-responsive element-binding protein. *J Biol Chem*, 269, 6405-6409.

Cairo, G. and Pietrangelo, A. (2000) Iron regulatory proteins in pathobiology. *Biochem J*, 352 Pt 2, 241-250.

Cairo, G., Tacchini, L. and Pietrangelo, A. (1998a) Lack of coordinate control of ferritin and transferrin receptor expression during rat liver regeneration. *Hepatology*, 28, 173-178.

Cairo, G., Tacchini, L., Pogliaghi, G., Anzon, E., Tomasi, A. and Bernelli-Zazzera, A. (1995) Induction of ferritin synthesis by oxidative stress. Transcriptional and post-transcriptional regulation by expansion of the "free" iron pool. *J Biol Chem*, 270, 700-703.

Cairo, G., Tacchini, L., Recalcati, S., Azzimonti, B., Minotti, G. and Bernelli-Zazzera, A. (1998b) Effect of reactive oxygen species on iron regulatory protein activity. *Ann N Y Acad Sci*, 851, 179-186.

Camaschella, C., Roetto, A., Cali, A., De Gobbi, M., Garozzo, G., Carella, M., Majorano, N., Totaro, A. and Gasparini, P. (2000) The gene TFR2 is mutated in a new type of haemochromatosis mapping to 7q22. *Nat Genet*, 25, 14-15.

Canonne-Hergaux, F., Gruenheid, S., Ponka, P. and Gros, P. (1999) Cellular and subcellular localization of the Nramp2 iron transporter in the intestinal brush border and regulation by dietary iron. *Blood*, 93, 4406-4417.

Canonne-Hergaux, F., Levy, J.E., Fleming, M.D., Montross, L.K., Andrews, N.C. and Gros, P. (2001) Expression of the DMT1 (NRAMP2/DCT1) iron transporter in mice with genetic iron overload disorders. *Blood*, 97, 1138-1140.

Carter, C.W., Jr. and Carter, C.W. (1979) Protein crystallization using incomplete factorial experiments. *J Biol Chem*, 254, 12219-12223.

Carter, C.W. and Sweet, R.M. (1997) *Macromolecular Crystallography, Part A. Methods Enzym.*, 276.

Cazzola, M., Bergamaschi, G., Tonon, L., Arbustini, E., Grasso, M., Vercesi, E., Barosi, G., Bianchi, P.E., Cairo, G. and Arosio, P. (1997) Hereditary hyperferritinemia-cataract syndrome: relationship between phenotypes and specific mutations in the iron-responsive element of ferritin light-chain mRNA. *Blood*, 90, 814-821.

Chabriere, E., Volbeda, A., Fontecilla-Camps, J.C., Roth, M. and Charon, M.H. (1999) Combination of methods used in the structure solution of pyruvate:ferredoxin oxidoreductase from two crystal forms. *Acta Crystallogr D Biol Crystallogr*, 55 (Pt 9), 1546-1554.

Chayen, N. (1997) A novel technique to control the rate of vapour diffusion, giving larger protein crystals. *Journal of Applied Crystallography*, 30, 198-202.

Chen, O.S., Hemenway, S. and Kaplan, J. (2002) Inhibition of Fe-S cluster biosynthesis decreases mitochondrial iron export: evidence that Yfh1p affects Fe-S cluster synthesis. *Proc Natl Acad Sci U S A*, 99, 12321-12326.

Chen, X.J., Wang, X., Kaufman, B.A. and Butow, R.A. (2005) Aconitase couples metabolic regulation to mitochondrial DNA maintenance. *Science*, 307, 714-717.

Cheng, Y., Zak, O., Aisen, P., Harrison, S.C. and Walz, T. (2004) Structure of the human transferrin receptor-transferrin complex. *Cell*, 116, 565-576.

Cleland, W.W. (2000) Low-barrier hydrogen bonds and enzymatic catalysis. *Arch Biochem Biophys*, 382, 1-5.

Collaborative Computational Project, N. (1994) The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst. D*, 50, 760-763.

Connor, J.R., Milward, E.A., Moalem, S., Sampietro, M., Boyer, P., Percy, M.E., Vergani, C., Scott, R.J. and Chorney, M. (2001) Is hemochromatosis a risk factor for Alzheimer's disease? *J Alzheimers Dis*, 3, 471-477.

Connor, J.R., Snyder, B.S., Arosio, P., Loeffler, D.A. and LeWitt, P. (1995) A quantitative analysis of isoferritins in select regions of aged, parkinsonian, and Alzheimer's diseased brains. *J Neurochem*, 65, 717-724.

Connor, J.R., Snyder, B.S., Beard, J.L., Fine, R.E. and Mufson, E.J. (1992) Regional distribution of iron and iron-regulatory proteins in the brain in aging and Alzheimer's disease. *J Neurosci Res*, 31, 327-335.

Cowtan, K.D. and Main, P. (1996) Phase combination and cross validation in iterated density-modification calculations. *Acta Cryst.*, D52, 43-48.

Cox, T.C., Bawden, M.J., Martin, A. and May, B.K. (1991) Human erythroid 5-aminolevulinate synthase: promoter analysis and identification of an iron-responsive element in the mRNA. *Embo J*, 10, 1891-1902.

Cromer, D.T. and Liberman, D. (1970) Relativistic Calculation of Anomalous Scattering Factors for X Rays. *The Journal of Chemical Physics*, 53, 1891-1898.

Crompton, D.E., Chinnery, P.F., Fey, C., Curtis, A.R., Morris, C.M., Kierstan, J., Burt, A., Young, F., Coulthard, A., Curtis, A., Ince, P.G., Bates, D., Jackson, M.J. and Burn, J. (2002) Neuroferritinopathy: a window on the role of iron in neurodegeneration. *Blood Cells Mol Dis*, 29, 522-531.

Crowther, R.A. and Blow, D.M. (1967) A method for positioning a known molecule in a unknown crystal structure. *Acta Cryst*, 23, 544-548.

Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y. and McPherson, A. (1994) Screening and optimization strategies for macromolecular crystal growth. *Acta Cryst.*, D50, 414-423.

Cunningham, L., Gruer, M.J. and Guest, J.R. (1997) Transcriptional regulation of the aconitase genes (*acnA* and *acnB*) of *Escherichia coli*. *Microbiology*, 143 (Pt 12), 3795-3805.

Curtis, A.R., Fey, C., Morris, C.M., Bindoff, L.A., Ince, P.G., Chinnery, P.F., Coulthard, A., Jackson, M.J., Jackson, A.P., McHale, D.P., Hay, D., Barker, W.A., Markham, A.F., Bates, D., Curtis, A. and Burn, J. (2001) Mutation in the gene encoding ferritin light polypeptide causes dominant adult-onset basal ganglia disease. *Nat Genet*, 28, 350-354.

D

Dandekar, T., Stripecke, R., Gray, N.K., Goossen, B., Constable, A., Johansson, H.E. and Hentze, M.W. (1991) Identification of a novel iron-responsive element in murine and human erythroid delta-aminolevulinic acid synthase mRNA. *Embo J*, 10, 1903-1909.

Dang, C.V. and Semenza, G.L. (1999) Oncogenic alterations of metabolism. *Trends Biochem Sci*, 24, 68-72.

Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G. and Sheldrick, G.M. (1999) Can anomalous signal of sulfur become a tool for solving protein crystal structures? *J Mol Biol*, 289, 83-92.

de la Fortelle, E. and Bricogne, G. (1997) Maximum-Likelihood Heavy-Atom Parameter Refinement for Multiple Isomorphous Replacement and Multiwavelength Anomalous Diffraction Methods. *Methods in Enzymology*, 276, 472-494.

DeRusso, P.A., Philpott, C.C., Iwai, K., Mostowski, H.S., Klausner, R.D. and Rouault, T.A. (1995) Expression of a constitutive mutant of iron regulatory protein 1 abolishes iron homeostasis in mammalian cells. *J Biol Chem*, 270, 15451-15454.

Dexter, D.T., Wells, F.R., Agid, F., Agid, Y., Lees, A.J., Jenner, P. and Marsden, C.D. (1987) Increased nigral iron content in postmortem parkinsonian brain. *Lancet*, 2, 1219-1220.

Diederichs, K. and Karplus, P.A. (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol*, 4, 269-275.

Donovan, A., Brownlie, A., Zhou, Y., Shepard, J., Pratt, S.J., Moynihan, J., Paw, B.H., Drejer, A., Barut, B., Zapata, A., Law, T.C., Brugnara, C., Lux, S.E., Pinkus, G.S., Pinkus, J.L., Kingsley, P.D., Palis, J., Fleming, M.D., Andrews, N.C. and Zon, L.I. (2000) Positional cloning of zebrafish ferroportin1 identifies a conserved vertebrate iron exporter. *Nature*, 403, 776-781.

Doublie, S. (1997) Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol*, 276, 523-530.

Drapier, J.C. (1997) Interplay between NO and [Fe-S] clusters: relevance to biological systems. *Methods*, 11, 319-329.

Drapier, J.C. and Bouton, C. (1996) Modulation by nitric oxide of metalloprotein regulatory activities. *Bioessays*, 18, 549-556.

Drapier, J.C., Hirling, H., Wietzerbin, J., Kaldy, P. and Kühn, L.C. (1993) Biosynthesis of nitric oxide activates iron regulatory factor in macrophages. *Embo J*, 12, 3643-3649.

Ducruix, A. and Giege, R. (1992) Crystallization of Nucleic Acids and Proteins - A Practical Approach. IRL, Oxford.

Duff, J.L.C., Breton, J.L.J., Butt, J.N., Armstrong, F.A. and Thomson, A.J. (1996) Novel Redox Chemistry of [3Fe-4S] Clusters: Electrochemical Characterization of the All-Fe(II) Form of the [3Fe-4S] Cluster Generated Reversibly in Various Proteins and Its Spectroscopic Investigation in *Sulfolobus acidocaldarius* Ferredoxin. *J. Am. Chem. Soc.*, 118, 8593-8603.

E

Eisenstein, R.S. (2000) Iron regulatory proteins and the molecular control of mammalian iron metabolism. *Annu Rev Nutr*, 20, 627-662.

Engh, R.A. and Huber, R. (1991) Accurate Bond and Angle Parameters for X-ray Protein Structure Refinement. *Acta Cryst.*, A47, 392-400.

Epsztejn, S., Glickstein, H., Picard, V., Slotki, I.N., Breuer, W., Beaumont, C. and Cabantchik, Z.I. (1999) H-ferritin subunit overexpression in erythroid cells reduces the oxidative stress response and induces multidrug resistance properties. *Blood*, 94, 3593-3603.

Evans, G. and Pettifer, R.F. (2001) CHOOCH: a program for deriving anomalous-scattering factors from X-ray fluorescence spectra. *J. Appl. Cryst.*, 34, 82-86.

Emery-Goodman, A., Hirling, H., Scarpellino, L., Henderson, B. and Kühn, L.C. (1993) Iron regulatory factor expressed from recombinant baculovirus: conversion between the RNA-binding apoprotein and Fe-S cluster containing aconitase. *Nucleic Acids Res*, 21, 1457-1461.

F

Faucheux, B.A., Bonnet, A.M., Agid, Y. and Hirsch, E.C. (1999) Blood vessels change in the mesencephalon of patients with Parkinson's disease. *Lancet*, 353, 981-982.

Faucheux, B.A., Martin, M.E., Beaumont, C., Hauw, J.J., Agid, Y. and Hirsch, E.C. (2003) Neuromelanin associated redox-active iron is increased in the substantia nigra of patients with Parkinson's disease. *J Neurochem*, 86, 1142-1148.

Faucheux, B.A., Nillesse, N., Damier, P., Spik, G., Mouatt-Prigent, A., Pierce, A., Leveugle, B., Kubis, N., Hauw, J.J., Agid, Y. and et al. (1995) Expression of lactoferrin receptors is increased in the mesencephalon of patients with Parkinson disease. *Proc Natl Acad Sci U S A*, 92, 9603-9607.

Ferguson, C.J., Wareing, M., Ward, D.T., Green, R., Smith, C.P. and Riccardi, D. (2001) Cellular localization of divalent metal transporter DMT-1 in rat kidney. *Am J Physiol Renal Physiol*, 280, F803-814.

Fillebeen, C., Caltagirone, A., Martelli, A., Moulis, J.M. and Pantopoulos, K. (2005) IRP1 Ser-711 is a phosphorylation site, critical for regulation of RNA-binding and aconitase activities. *Biochem J*.

Fillebeen, C., Chahine, D., Caltagirone, A., Segal, P. and Pantopoulos, K. (2003) A phosphomimetic mutation at Ser-138 renders iron regulatory protein 1 sensitive to iron-dependent degradation. *Mol Cell Biol*, 23, 6973-6981.

Finch, C. (1994) Regulators of iron balance in humans. *Blood*, 84, 1697-1702.

Fleming, M.D., Romano, M.A., Su, M.A., Garrick, L.M., Garrick, M.D. and Andrews, N.C. (1998) Nramp2 is mutated in the anemic Belgrade (b) rat: evidence of a role for Nramp2 in endosomal iron transport. *Proc Natl Acad Sci U S A*, 95, 1148-1153.

Fleming, M.D., Trenor, C.C., 3rd, Su, M.A., Foernzler, D., Beier, D.R., Dietrich, W.F. and Andrews, N.C. (1997) Microcytic anaemia mice have a mutation in Nramp2, a candidate iron transporter gene. *Nat Genet*, 16, 383-386.

Fleming, R.E., Migas, M.C., Holden, C.C., Waheed, A., Britton, R.S., Tomatsu, S., Bacon, B.R. and Sly, W.S. (2000) Transferrin receptor 2: continued expression in mouse liver in the face of iron overload and in hereditary hemochromatosis. *Proc Natl Acad Sci U S A*, 97, 2214-2219.

G

Gaudu, P. and Weiss, B. (1996) SoxR, a [2Fe-2S] transcription factor, is active only in its oxidized form. *Proc Natl Acad Sci U S A*, 93, 10094-10098.

Gegout, V., Schlegl, J., Schlager, B., Hentze, M.W., Reinbolt, J., Ehresmann, B., Ehresmann, C. and Romby, P. (1999) Ligand-induced structural alterations in human iron regulatory protein-1 revealed by protein footprinting. *J Biol Chem*, 274, 15052-15058.

- Gehring, N.H., Hentze, M.W. and Pantopoulos, K. (1999) Inactivation of both RNA binding and aconitase activities of iron regulatory protein-1 by quinone-induced oxidative stress. *J Biol Chem*, 274, 6219-6225.
- Gerber, J. and M Lill, R. (2002) Biogenesis of iron-sulfur protéins ineukaryotes: components, mechanism and pathology. *Mitochondrion*, 2, 71-86.
- Gerber, J., Muhlenhoff, U. and Lill, R. (2003) An interaction between frataxin and Isu1/Nfs1 that is crucial for Fe/S cluster synthesis on Isu1. *EMBO Rep*, 4, 906-911.
- Gerlach, M., Ben-Shachar, D., Riederer, P. and Youdim, M.B. (1994) Altered brain metabolism of iron as a cause of neurodegenerative diseases? *J Neurochem*, 63, 793-807.
- Goessling, L.S., Mascotti, D.P. and Thach, R.E. (1998) Involvement of heme in the degradation of iron-regulatory protein 2. *J Biol Chem*, 273, 12555-12557.
- Golbeck, J.H. and Bryant, D.A. (1991) Photosystem I. *Curr Top Bioenerg*, 16, 3-177.
- Goodsell, D.S., Lauble, H., Stout, C.D. and Olson, A.J. (1993) Automated docking in crystallography: analysis of the substrates of aconitase. *Proteins*, 17, 1-10.
- Gordon, N. (2000) Friedreich's ataxia and iron metabolism. *Brain Dev*, 22, 465-468.
- Gotz, M.E., Double, K., Gerlach, M., Youdim, M.B. and Riederer, P. (2004) The relevance of iron in the pathogenesis of Parkinson's disease. *Ann N Y Acad Sci*, 1012, 193-208.
- Grandoni, J.A., Switzer, R.L., Makaroff, C.A. and Zalkin, H. (1989) Evidence that the iron-sulfur cluster of *Bacillus subtilis* glutamine phosphoribosylpyrophosphate amidotransferase determines stability of the enzyme to degradation in vivo. *J Biol Chem*, 264, 6058-6064.
- Gray, N.K., Pantopoulos, K., Dandekar, T., Ackrell, B.A. and Hentze, M.W. (1996) Translational regulation of mammalian and *Drosophila* citric acid cycle enzymes via iron-responsive elements. *Proc Natl Acad Sci U S A*, 93, 4925-4930.
- Green, D.W., Ingram, V.M. and Perutz, M.F. (1954) The structure oh haemoglobin IV. Sign of determination by isomorphous replacement method. *Proc. R. Soc. London SER, A* 225, 287-307.
- Griffiths, P.D., Dobson, B.R., Jones, G.R. and Clarke, D.T. (1999) Iron in the basal ganglia in Parkinson's disease. An in vitro study using extended X-ray absorption fine structure and cryo-electron microscopy. *Brain*, 122 (Pt 4), 667-673.
- Gruer, M.J., Artymiuk, P.J. and Guest, J.R. (1997) The aconitase family: three structural variations on a common theme. *Trends Biochem Sci*, 22, 3-6.
- Gunshin, H., Mackenzie, B., Berger, U.V., Gunshin, Y., Romero, M.F., Boron, W.F., Nussberger, S., Gollan, J.L. and Hediger, M.A. (1997) Cloning and characterization of a mammalian proton-coupled metal-ion transporter. *Nature*, 388, 482-488.

H

- Haile, D.J., Rouault, T.A., Harford, J.B., Kennedy, M.C., Blondin, G.A., Beinert, H. and Klausner, R.D. (1992a) Cellular regulation of the iron-responsive element binding protein: disassembly of the cubane iron-sulfur cluster results in high-affinity RNA binding. *Proc Natl Acad Sci U S A*, 89, 11735-11739.
- Haile, D.J., Rouault, T.A., Tang, C.K., Chin, J., Harford, J.B. and Klausner, R.D. (1992b) Reciprocal control of RNA-binding and aconitase activity in the regulation of the iron-responsive element binding protein: role of the iron-sulfur cluster. *Proc Natl Acad Sci U S A*, 89, 7536-7540.
- Hallgren, B. and Sourander, P. (1958) The effect of age on the non-haemin iron in the human brain. *J Neurochem*, 3, 41-51.
- Hanson, E.S., Foot, L.M. and Leibold, E.A. (1999) Hypoxia post-translationally activates iron-regulatory protein 2. *J Biol Chem*, 274, 5047-5052.
- Hanson, E.S. and Leibold, E.A. (1998) Regulation of iron regulatory protein 1 during hypoxia and hypoxia/reoxygenation. *J Biol Chem*, 273, 7588-7593.
- Hanson, E.S., Rawlins, M.L. and Leibold, E.A. (2003) Oxygen and iron regulation of iron regulatory protein 2. *J Biol Chem*, 278, 40337-40342.
- Harada, K. and Frankel, A.D. (1995) Identification of two novel arginine binding DNAs. *Embo J*, 14, 5798-5811.
- Harris, Z.L., Durley, A.P., Man, T.K. and Gitlin, J.D. (1999) Targeted gene disruption reveals an essential role for ceruloplasmin in cellular iron efflux. *Proc Natl Acad Sci U S A*, 96, 10812-10817.
- Harris, Z.L., Takahashi, Y., Miyajima, H., Serizawa, M., MacGillivray, R.T. and Gitlin, J.D. (1995) Acéruloplasminemia: molecular characterization of this disorder of iron metabolism. *Proc Natl Acad Sci U S A*, 92, 2539-2543.
- Harrison, P.M. (1977) Ferritin: an iron-storage molecule. *Semin Hematol*, 14, 55-70.
- Hashimoto, M., Takeda, A., Hsu, L.J., Takenouchi, T. and Masliah, E. (1999) Role of cytochrome c as a stimulator of alpha-synuclein aggregation in Lewy body disease. *J Biol Chem*, 274, 28849-28852.
- Hauptman, H.A. (1997) Shake-and-bake: an algorithm for automatic solution ab initio of crystal structures. *Methods Enzym.*, 277, 3-13.
- Hayflick, S.J. (2003) Unraveling the Hallervorden-Spatz syndrome: pantothenate kinase-associated neurodegeneration is the name. *Curr Opin Pediatr*, 15, 572-577.
- Henderson, B.R. and Kühn, L.C. (1995) Differential modulation of the RNA-binding proteins IRP-1 and IRP-2 in response to iron. IRP-2 inactivation requires translation of another protein. *J Biol Chem*, 270, 20509-20515.

- Henderson, B.R., Menotti, E., Bonnard, C. and Kühn, L.C. (1994) Optimal sequence and structure of iron-responsive elements. Selection of RNA stem-loops with high affinity for iron regulatory factor. *J Biol Chem*, 269, 17481-17489.
- Henderson, B.R., Menotti, E. and Kühn, L.C. (1996) Iron regulatory proteins 1 and 2 bind distinct sets of RNA target sequences. *J Biol Chem*, 271, 4900-4908.
- Hendrickson, W.A., Horton, J.R. and LeMaster, D.M. (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *Embo J*, 9, 1665-1672.
- Hendrickson, W.A. and Lattman, E.E. (1970) Representation of phase probability distributions for simplified combination of independent phase information. *Acta Cryst.*, B26, 136-143.
- Hendrickson, W.A. and Ogata, C.M. (1997) Phase determination from multiwavelength anomalous diffraction measurements. *Methods in enzymology*, 276, 494-523.
- Hendrickson, W.A. and Teeter, M.M. (1981) Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulphur. *Nature*, 290, 107 - 113.
- Hentze, M.W. and Kühn, L.C. (1996) Molecular control of vertebrate iron metabolism: mRNA-based regulatory circuits operated by iron, nitric oxide, and oxidative stress. *Proc Natl Acad Sci U S A*, 93, 8175-8182.
- Hentze, M.W., Muckenthaler, M.U. and Andrews, N.C. (2004) Balancing acts: molecular control of mammalian iron metabolism. *Cell*, 117, 285-297.
- Hidalgo, E., Bollinger, J.M., Jr., Bradley, T.M., Walsh, C.T. and Dimple, B. (1995) Binuclear [2Fe-2S] clusters in the Escherichia coli SoxR protein and role of the metal centers in transcription. *J Biol Chem*, 270, 20908-20914.
- Hirling, H., Henderson, B.R. and Kühn, L.C. (1994) Mutational analysis of the [4Fe-4S]-cluster converting iron regulatory factor from its RNA-binding form to cytoplasmic aconitase. *Embo J*, 13, 453-461.
- Hirose, W., Ikematsu, K. and Tsuda, R. (2003) Age-associated increases in heme oxygenase-1 and ferritin immunoreactivity in the autopsied brain. *Leg Med (Tokyo)*, 5 Suppl 1, S360-366.
- Hirsch, E.C., Brandel, J.P., Galle, P., Javoy-Agid, F. and Agid, Y. (1991) Iron and aluminum increase in the substantia nigra of patients with Parkinson's disease: an X-ray microanalysis. *J Neurochem*, 56, 446-451.
- Howard, J.B. and Rees, D.C. (1996) Structural Basis of Biological Nitrogen Fixation. *Chem Rev*, 96, 2965-2982.
- Huang, X., Cuajungco, M.P., Atwood, C.S., Moir, R.D., Tanzi, R.E. and Bush, A.I. (2000) Alzheimer's disease, beta-amyloid protein and zinc. *J Nutr*, 130, 1488S-1492S.

Hubert, N. and Hentze, M.W. (2002) Previously uncharacterized isoforms of divalent metal transporter (DMT)-1: implications for regulation and cellular function. *Proc Natl Acad Sci U S A*, 99, 12345-12350.

Hussain, R.I., Ballard, C.G., Edwardson, J.A. and Morris, C.M. (2002) Transferrin gene polymorphism in Alzheimer's disease and dementia with Lewy bodies in humans. *Neurosci Lett*, 317, 13-16.

I

Islam, S.A., Carvin, D., Sternberg, M.J.E. and Blundell, T.L. (1998) HAD, a Data Bank of Heavy-Atom Binding Sites in Protein Crystals: a Resource for Use in Multiple Isomorphous Replacement and Anomalous Scattering. *Acta Cryst.*, D54, 1199-1206.

Iwai, K., Drake, S.K., Wehr, N.B., Weissman, A.M., LaVaute, T., Minato, N., Klausner, R.D., Levine, R.L. and Rouault, T.A. (1998) Iron-dependent oxidation, ubiquitination, and degradation of iron regulatory protein 2: implications for degradation of oxidized proteins. *Proc Natl Acad Sci U S A*, 95, 4924-4928.

Iwai, K., Klausner, R.D. and Rouault, T.A. (1995) Requirements for iron-regulated degradation of the RNA binding protein, iron regulatory protein 2. *Embo J*, 14, 5350-5357.

Iwasaki, T., Imai, T., Urushiyama, A. and Oshima, T. (1996) Redox-linked ionization of sulredoxin, an archaeal Rieske-type [2Fe-2S] protein from *Sulfolobus* sp. strain 7. *J Biol Chem*, 271, 27659-27663.

J

Jancarik, J. and Kim, S.H. (1991) Sparse matrix sampling: a screening method for crystallization of proteins. *Journal of Applied Crystallography*, 24, 409-411.

Jellinger, K., Paulus, W., Grundke-Iqbal, I., Riederer, P. and Youdim, M.B. (1990) Brain iron and ferritin in Parkinson's and Alzheimer's diseases. *J Neural Transm Park Dis Dement Sect*, 2, 327-340.

Jeong, J., Rouault, T.A. and Levine, R.L. (2004) Identification of a heme-sensing domain in iron regulatory protein 2. *J Biol Chem*, 279, 45450-45454.

Jeong, S.Y. and David, S. (2003) Glycosylphosphatidylinositol-anchored ceruloplasmin is required for iron efflux from cells in the central nervous system. *J Biol Chem*, 278, 27144-27148.

Johnson, M.K. (1998) Iron-sulfur proteins: new roles for old clusters. *Curr Opin Chem Biol*, 2, 173-181.

Jones, S., Daley, D.T., Luscombe, N.M., Berman, H.M. and Thornton, J.M. (2001) Protein-RNA interactions: a structural analysis. *Nucleic Acids Res*, 29, 943-954.

K

- Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.*, 26, 795-800.
- Kabsch, W. and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22, 2577-2637.
- Kaldy, P., Menotti, E., Moret, R. and Kühn, L.C. (1999) Identification of RNA-binding surfaces in iron regulatory protein-1. *Embo J*, 18, 6073-6083.
- Kang, D.K., Jeong, J., Drake, S.K., Wehr, N.B., Rouault, T.A. and Levine, R.L. (2003) Iron regulatory protein 2 as iron sensor. Iron-dependent oxidative modification of cysteine. *J Biol Chem*, 278, 14857-14864.
- Kawabata, H., Germain, R.S., Vuong, P.T., Nakamaki, T., Said, J.W. and Koeffler, H.P. (2000) Transferrin receptor 2- α supports cell growth both in iron-chelated cultured cells and in vivo. *J Biol Chem*, 275, 16618-16625.
- Kawabata, H., Yang, R., Hirama, T., Vuong, P.T., Kawano, S., Gombart, A.F. and Koeffler, H.P. (1999) Molecular cloning of transferrin receptor 2. A new member of the transferrin receptor-like family. *J Biol Chem*, 274, 20826-20832.
- Ke, Y. and Ming Qian, Z. (2003) Iron misregulation in the brain: a primary cause of neurodegenerative disorders. *Lancet Neurol*, 2, 246-253.
- Ke, Y., Wu, J., Leibold, E.A., Walden, W.E. and Theil, E.C. (1998) Loops and bulge/loops in iron-responsive element isoforms influence iron regulatory protein binding. Fine-tuning of mRNA regulation? *J Biol Chem*, 273, 23637-23640.
- Kennedy, M.C., Antholine, W.E. and Beinert, H. (1997) An EPR investigation of the products of the reaction of cytosolic and mitochondrial aconitases with nitric oxide. *J Biol Chem*, 272, 20340-20347.
- Kennedy, M.C., Mende-Mueller, L., Blondin, G.A. and Beinert, H. (1992) Purification and characterization of cytosolic aconitase from beef liver and its relationship to the iron-responsive element binding protein. *Proc Natl Acad Sci U S A*, 89, 11730-11734.
- Khoroshilova, N., Beinert, H. and Kiley, P.J. (1995) Association of a polynuclear iron-sulfur center with a mutant FNR protein enhances DNA binding. *Proc Natl Acad Sci U S A*, 92, 2499-2503.
- Kim, H.Y., Klausner, R.D. and Rouault, T.A. (1995) Translational repressor activity is equivalent and is quantitatively predicted by in vitro RNA binding for two iron-responsive element-binding proteins, IRP1 and IRP2. *J Biol Chem*, 270, 4983-4986.
- Kim, S. and Ponka, P. (1999) Control of transferrin receptor expression via nitric oxide-mediated modulation of iron-regulatory protein 2. *J Biol Chem*, 274, 33035-33042.

Kim, S. and Ponka, P. (2000) Effects of interferon-gamma and lipopolysaccharide on macrophage iron metabolism are mediated by nitric oxide-induced degradation of iron regulatory protein 2. *J Biol Chem*, 275, 6220-6226.

Kim, S. and Ponka, P. (2002) Nitrogen monoxide-mediated control of ferritin synthesis: implications for macrophage iron homeostasis. *Proc Natl Acad Sci U S A*, 99, 12214-12219.

Kim, S.Y., Tak, J.K. and Park, J.W. (2004) Inactivation of NADP(+)-dependent isocitrate dehydrogenase by singlet oxygen derived from photoactivated rose bengal. *Biochimie*, 86, 501-507.

Kleywegt, G.J. and Brunger, A.T. (1996) Checking your imagination: applications of the free R value. *Structure*, 4, 897-904.

Kohler, S.A., Henderson, B.R. and Kühn, L.C. (1995) Succinate dehydrogenase b mRNA of *Drosophila melanogaster* has a functional iron-responsive element in its 5'-untranslated region. *J Biol Chem*, 270, 30781-30786.

Kohler, S.A., Menotti, E. and Kühn, L.C. (1999) Molecular cloning of mouse glycolate oxidase. High evolutionary conservation and presence of an iron-responsive element-like sequence in the mRNA. *J Biol Chem*, 274, 2401-2407.

Kozyraki, R., Fyfe, J., Verroust, P.J., Jacobsen, C., Dautry-Varsat, A., Gburek, J., Willnow, T.E., Christensen, E.I. and Moestrup, S.K. (2001) Megalin-dependent cubilin-mediated endocytosis is a major pathway for the apical uptake of transferrin in polarized epithelia. *Proc Natl Acad Sci U S A*, 98, 12491-12496.

Krause, A., Neitz, S., Magert, H.J., Schulz, A., Forssmann, W.G., Schulz-Knappe, P. and Adermann, K. (2000) LEAP-1, a novel highly disulfide-bonded human peptide, exhibits antimicrobial activity. *FEBS Lett*, 480, 147-150.

Kristiansen, M., Graversen, J.H., Jacobsen, C., Sonne, O., Hoffman, H.J., Law, S.K. and Moestrup, S.K. (2001) Identification of the haemoglobin scavenger receptor. *Nature*, 409, 198-201.

Kronig, R. and Kramers, H.A. (1928) *Z. Phys.*, 48, 174.

Kwak, E.L., Larochelle, D.A., Beaumont, C., Torti, S.V. and Torti, F.M. (1995) Role for NF-kappa B in the regulation of ferritin H by tumor necrosis factor-alpha. *J Biol Chem*, 270, 15285-15293.

L

Langston, J.W., Forno, L.S., Tetrad, J., Reeves, A.G., Kaplan, J.A. and Karluk, D. (1999) Evidence of active nerve cell degeneration in the substantia nigra of humans years after 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine exposure. *Ann Neurol*, 46, 598-605.

Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, 26, 283-291.

Lauble, H., Kennedy, M.C., Beinert, H. and Stout, C.D. (1992) Crystal structures of aconitase with isocitrate and nitroisocitrate bound. *Biochemistry*, 31, 2735-2748.

Lauble, H., Kennedy, M.C., Beinert, H. and Stout, C.D. (1994) Crystal structures of aconitase with trans-aconitate and nitrocitrate bound. *J Mol Biol*, 237, 437-451.

LaVaute, T., Smith, S., Cooperman, S., Iwai, K., Land, W., Meyron-Holtz, E., Drake, S.K., Miller, G., Abu-Asab, M., Tsokos, M., Switzer, R., 3rd, Grinberg, A., Love, P., Tresser, N. and Rouault, T.A. (2001) Targeted deletion of the gene encoding iron regulatory protein-2 causes misregulation of iron metabolism and neurodegenerative disease in mice. *Nat Genet*, 27, 209-214.

Leenders, K.L.e.a. Verapamil uptake in mesecephalon of Parkinson disease. *Ann. Neurol.* (in the press).

Leveugle, B., Faucheux, B.A., Bouras, C., Nillesse, N., Spik, G., Hirsch, E.C., Agid, Y. and Hof, P.R. (1996) Cellular distribution of the iron-binding protein lactotransferrin in the mesencephalon of Parkinson's disease cases. *Acta Neuropathol (Berl)*, 91, 566-572.

Levi, S., Corsi, B., Bosisio, M., Invernizzi, R., Volz, A., Sanford, D., Arosio, P. and Drysdale, J. (2001) A human mitochondrial ferritin encoded by an intronless gene. *J Biol Chem*, 276, 24437-24440.

Levine, S., Connor, J. and Schipper, H.M. (2004) Redox-active metals in neurological disorders. *Ann. NY Acad. Sci.*, 1012.

Levy, J.E., Jin, O., Fujiwara, Y., Kuo, F. and Andrews, N.C. (1999) Transferrin receptor is necessary for development of erythrocytes and the nervous system. *Nat Genet*, 21, 396-399.

Lobreaux, S., Thoiron, S. and Briat, J.F. (1995) Induction of ferritin synthesis in maize leaves by an iron-mediated oxidative stress. *Plant J.*, 8, 443-449.

Lok, C.N. and Ponka, P. (1999) Identification of a hypoxia response element in the transferrin receptor gene. *J Biol Chem*, 274, 24147-24152.

M

Mantyh, P.W., Ghilardi, J.R., Rogers, S., DeMaster, E., Allen, C.J., Stimson, E.R. and Maggio, J.E. (1993) Aluminum, iron, and zinc ions promote aggregation of physiological concentrations of beta-amyloid peptide. *J Neurochem*, 61, 1171-1174.

Martins, E.A., Robalinho, R.L. and Meneghini, R. (1995) Oxidative stress induces activation of a cytosolic protein responsible for control of iron uptake. *Arch Biochem Biophys*, 316, 128-134.

Matthews, B.W. (1968) Solvent content of protein crystals. *J Mol Biol*, 33, 491-497.

McKie, A.T., Barrow, D., Latunde-Dada, G.O., Rolfs, A., Sager, G., Mudaly, E., Mudaly, M., Richardson, C., Barlow, D., Bomford, A., Peters, T.J., Raja, K.B., Shirali, S., Hediger, M.A., Farzaneh, F. and Simpson, R.J. (2001) An iron-regulated ferric reductase associated with the absorption of dietary iron. *Science*, 291, 1755-1759.

McKie, A.T., Marciani, P., Rolfs, A., Brennan, K., Wehr, K., Barrow, D., Miret, S., Bomford, A., Peters, T.J., Farzaneh, F., Hediger, M.A., Hentze, M.W. and Simpson, R.J. (2000) A novel

duodenal iron-regulated transporter, IREG1, implicated in the basolateral transfer of iron to the circulation. *Mol Cell*, 5, 299-309.

McPherson, A. (1989) Preparation and analysis of protein crystals. Robert E. Krieger Publishing Company, Second Edition.

McPherson, A. (1992) Two approaches to the rapid screening of crystallization conditions. *J Cryst Growth*, 122, 161-167.

McPherson, A., Malkin, A.J. and Kuznetsov, Y.G. (1995) The science of macromolecular crystallization. *Structure*, 3, 759-768.

Menotti, E., Henderson, B.R. and Kühn, L.C. (1998) Translational regulation of mRNAs with distinct IRE sequences by iron regulatory proteins 1 and 2. *J Biol Chem*, 273, 1821-1824.

Meyron-Holtz, E.G., Ghosh, M.C., Iwai, K., LaVaute, T., Brazzolotto, X., Berger, U.V., Land, W., Ollivierre-Wilson, H., Grinberg, A., Love, P. and Rouault, T.A. (2004) Genetic ablations of iron regulatory proteins 1 and 2 reveal why iron regulatory protein 2 dominates iron homeostasis. *Embo J*, 23, 386-395.

Meyron-Holtz, E.G., Vaisman, B., Cabantchik, Z.I., Fibach, E., Rouault, T.A., Hershko, C. and Konijn, A.M. (1999) Regulation of intracellular iron metabolism in human erythroid precursors by internalized extracellular ferritin. *Blood*, 94, 3205-3211.

Mikulits, W., Schranzhofer, M., Bauer, A., Dolznig, H., Lobmayr, L., Infante, A.A., Beug, H. and Mullner, E.W. (1999) Impaired ferritin mRNA translation in primary erythroid progenitors: shift to iron-dependent regulation by the v-ErbA oncoprotein. *Blood*, 94, 4321-4332.

Miyajama, H.e.a. (1997) The use of deferrioxamine in the treatment of acéruloplasminemia. *Ann. Neurol.*, 41, 404-407.

Moalem, S., Percy, M.E., Andrews, D.F., Kruck, T.P., Wong, S., Dalton, A.J., Mehta, P., Fedor, B. and Warren, A.C. (2000) Are hereditary hemochromatosis mutations involved in Alzheimer disease? *Am J Med Genet*, 93, 58-66.

Muckenthaler, M., Gray, N.K. and Hentze, M.W. (1998) IRP-1 binding to ferritin mRNA prevents the recruitment of the small ribosomal subunit by the cap-binding complex eIF4F. *Mol Cell*, 2, 383-388.

Muckenthaler, M., Roy, C.N., Custodio, A.O., Minana, B., deGraaf, J., Montross, L.K., Andrews, N.C. and Hentze, M.W. (2003) Regulatory defects in liver and intestine implicate abnormal hepcidin and Cybrd1 expression in mouse hemochromatosis. *Nat Genet*, 34, 102-107.

Mueller, S. and Pantopoulos, K. (2002) Activation of iron regulatory protein-1 by oxidative stress. *Methods Enzymol*, 348, 324-337.

Mueller, S., Pantopoulos, K., Hubner, C.A., Stremmel, W. and Hentze, M.W. (2001) IRP1 activation by extracellular oxidative stress in the perfused rat liver. *J Biol Chem*, 276, 23192-23196.

Muhlenhoff, U., Gerber, J., Richhardt, N. and Lill, R. (2003) Components involved in assembly and dislocation of iron-sulfur clusters on the scaffold protein Isu1p. *Embo J*, 22, 4815-4825.

Muhlenhoff, U., Richhardt, N., Ristow, M., Kispal, G. and Lill, R. (2002) The yeast frataxin homolog Yfh1p plays a specific role in the maturation of cellular Fe/S proteins. *Hum Mol Genet*, 11, 2025-2036.

Mukhopadhyay, C.K., Attieh, Z.K. and Fox, P.L. (1998) Role of ceruloplasmin in cellular iron uptake. *Science*, 279, 714-717.

Mulero, V. and Brock, J.H. (1999) Regulation of iron metabolism in murine J774 macrophages: role of nitric oxide-dependent and -independent pathways following activation with gamma interferon and lipopolysaccharide. *Blood*, 94, 2383-2389.

Münch, G.e.a. (2000) Crosslinking of α -synuclein by advanced glycation endproducts — an early pathophysiological step in Lewy body formation? *J. Chem. Neuroanat.*, 253–257.

Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst D*, 53, 240-253.

Mwanjewe, J. and Grover, A.K. (2004) Role of transient receptor potential canonical 6 (TRPC6) in non-transferrin-bound iron uptake in neuronal phenotype PC12 cells. *Biochem J*, 378, 975-982.

N

Narahari, J., Ma, R., Wang, M. and Walden, W.E. (2000) The aconitase function of iron regulatory protein 1. Genetic studies in yeast implicate its role in iron-mediated redox regulation. *J Biol Chem*, 275, 16227-16234.

Navaza, J. (1994) AMoRe: an automated package for molecular replacement. *Acta Cryst.*, A50, 157-163.

Nemeth, E., Rivera, S., Gabayan, V., Keller, C., Taudorf, S., Pedersen, B.K. and Ganz, T. (2004) IL-6 mediates hypoferremia of inflammation by inducing the synthesis of the iron regulatory hormone hepcidin. *J Clin Invest*, 113, 1271-1276.

Nemeth, E., Valore, E.V., Territo, M., Schiller, G., Lichtenstein, A. and Ganz, T. (2003) Hepcidin, a putative mediator of anemia of inflammation, is a type II acute-phase protein. *Blood*, 101, 2461-2463.

Nicolas, G., Bennoun, M., Devaux, I., Beaumont, C., Grandchamp, B., Kahn, A. and Vaulont, S. (2001) Lack of hepcidin gene expression and severe tissue iron overload in upstream stimulatory factor 2 (USF2) knockout mice. *Proc Natl Acad Sci U S A*, 98, 8780-8785.

Nicolas, G., Bennoun, M., Porteu, A., Mativet, S., Beaumont, C., Grandchamp, B., Sirito, M., Sawadogo, M., Kahn, A. and Vaulont, S. (2002a) Severe iron deficiency anemia in transgenic mice expressing liver hepcidin. *Proc Natl Acad Sci U S A*, 99, 4596-4601.

Nicolas, G., Chauvet, C., Viatte, L., Danan, J.L., Bigard, X., Devaux, I., Beaumont, C., Kahn, A. and Vaulont, S. (2002b) The gene encoding the iron regulatory peptide hepcidin is regulated by anemia, hypoxia, and inflammation. *J Clin Invest*, 110, 1037-1044.

Nissen, P., Kjeldgaard, M., Thirup, S., Polekhina, G., Reshetnikova, L., Clark, B.F. and Nyborg, J. (1995) Crystal structure of the ternary complex of Phe-tRNAPhe, EF-Tu, and a GTP analog. *Science*, 270, 1464-1472.

Nordman, C.E. (1966) *Trans. Am. Cryst. Assoc.*, 2, 29-38.

O

Oliveira, L., Bouton, C. and Drapier, J.C. (1999) Thioredoxin activation of iron regulatory proteins. Redox regulation of RNA binding after exposure to nitric oxide. *J Biol Chem*, 274, 516-521.

Oliveira, L. and Drapier, J.C. (2000) Down-regulation of iron regulatory protein 1 gene expression by nitric oxide. *Proc Natl Acad Sci U S A*, 97, 6550-6555.

Ostrerova-Golts, N., Petrucelli, L., Hardy, J., Lee, J.M., Farer, M. and Woloizin, B. (2000) The A53T alpha-synuclein mutation increases iron-dependent aggregation and toxicity. *J Neurosci*, 20, 6048-6054.

Otwinowski, Z. (1991) Maximum likelihood refinement of heavy parameters. Isomorphous replacement and anomalous scattering. In *proceeding of the CCP4 study weekend*, 80-86.

Oudit, G.Y., Sun, H., Trivieri, M.G., Koch, S.E., Dawood, F., Ackerley, C., Yazdanpanah, M., Wilson, G.J., Schwartz, A., Liu, P.P. and Backx, P.H. (2003) L-type Ca²⁺ channels provide a major pathway for iron entry into cardiomyocytes in iron-overload cardiomyopathy. *Nat Med*, 9, 1187-1194.

P

Pandolfo, M. (2002) Iron metabolism and mitochondrial abnormalities in Friedreich ataxia. *Blood Cells Mol Dis*, 29, 536-547; discussion 548-552.

Pantopoulos, K. and Hentze, M.W. (1995) Rapid responses to oxidative stress mediated by iron regulatory protein. *Embo J*, 14, 2917-2924.

Pantopoulos, K. and Hentze, M.W. (1998) Activation of iron regulatory protein-1 by oxidative stress in vitro. *Proc Natl Acad Sci U S A*, 95, 10559-10563.

Pantopoulos, K. and Hentze, M.W. (2000) Nitric oxide, oxygen radicals, and iron metabolism. Academic Press. San Diego, 293-313.

Pantopoulos, K., Mueller, S., Atzberger, A., Ansorge, W., Stremmel, W. and Hentze, M.W. (1997) Differences in the regulation of iron regulatory protein-1 (IRP-1) by extra- and intracellular oxidative stress. *J Biol Chem*, 272, 9802-9808.

Pantopoulos, K., Weiss, G. and Hentze, M.W. (1996) Nitric oxide and oxidative stress (H₂O₂) control mammalian iron metabolism by different pathways. *Mol Cell Biol*, 16, 3781-3788.

Paraskeva, E. and Hentze, M.W. (1996) Iron-sulphur clusters as genetic regulatory switches: the bifunctional iron regulatory protein-1. *FEBS Lett*, 389, 40-43.

Park, C.H., Valore, E.V., Waring, A.J. and Ganz, T. (2001) Hepcidin, a urinary antimicrobial peptide synthesized in the liver. *J Biol Chem*, 276, 7806-7810.

Patel, P.I. and Isaya, G. (2001) Friedreich ataxia: from GAA triplet-repeat expansion to frataxin deficiency. *Am J Hum Genet*, 69, 15-24.

Perry, G., Sayre, L.M., Atwood, C.S., Castellani, R.J., Cash, A.D., Rotkamp, C.A. and Smith, M.A. (2002) The role of iron and copper in the aetiology of neurodegenerative disorders: therapeutic implications. *CNS Drugs*, 16, 339-352.

Philpott, C.C., Klausner, R.D. and Rouault, T.A. (1994) The bifunctional iron-responsive element binding protein/cytosolic aconitase: the role of active-site residues in ligand binding and regulation. *Proc Natl Acad Sci U S A*, 91, 7321-7325.

Pigeon, C., Ilyin, G., Courselaud, B., Leroyer, P., Turlin, B., Brissot, P. and Loreal, O. (2001) A new mouse liver-specific gene, encoding a protein homologous to human antimicrobial peptide hepcidin, is overexpressed during iron overload. *J Biol Chem*, 276, 7811-7819.

Pitula, J.S., Deck, K.M., Clarke, S.L., Anderson, S.A., Vasanthakumar, A. and Eisenstein, R.S. (2004) Selective inhibition of the citrate-to-isocitrate reaction of cytosolic aconitase by phosphomimetic mutation of serine-711. *Proc Natl Acad Sci U S A*, 101, 10907-10912.

Ponka, P. (1997) Tissue-specific regulation of iron metabolism and heme synthesis: distinct control mechanisms in erythroid cells. *Blood*, 89, 1-25.

Puglisi, J.D., Tan, R., Calnan, B.J., Frankel, A.D. and Williamson, J.R. (1992) Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science*, 257, 76-80.

Pulliam, J.F., Jennings, C.D., Kryscio, R.J., Davis, D.G., Wilson, D., Montine, T.J., Schmitt, F.A. and Markesbery, W.R. (2003) Association of HFE mutations with neurodegeneration and oxidative stress in Alzheimer's disease and correlation with APOE. *Am J Med Genet B Neuropsychiatr Genet*, 119, 48-53.

Q

Qi, Y., Jamindar, T.M. and Dawson, G. (1995) Hypoxia alters iron homeostasis and induces ferritin synthesis in oligodendrocytes. *J Neurochem*, 64, 2458-2464.

Qian, Z.M. and Wang, Q. (1998) Expression of iron transport proteins and excessive iron accumulation in the brain in neurodegenerative disorders. *Brain Res Brain Res Rev*, 27, 257-267.

R

- Ramachandran, G.N. and Raman, S. (1956) A new method for the structure analysis of non-centrosymmetric crystals. *Current Science*, 11, 348-351.
- Ramakrishnan, C. and Ramachandran, G.N. (1965) Stereochemical criteria for polypeptide and protein chain conformation. *Biophys. J.*, 5, 909-933.
- Ramakrishnan, V. and Biou, V. (1996) Treatment of multiwavelength anomalous diffraction data as a special case of multiple isomorphous replacement. *Methods in enzymology*, 276, 538-557.
- Read, R.J. (2001) Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Cryst.*, D57, 1373-1382.
- Recalcati, S., Taramelli, D., Conte, D. and Cairo, G. (1998) Nitric oxide-mediated induction of ferritin synthesis in J774 macrophages by inflammatory cytokines: role of selective iron regulatory protein-2 downregulation. *Blood*, 91, 1059-1066.
- Richardson, J.S. (1981) The anatomy and taxonomy of protein structure. *Adv Protein Chem*, 34, 167-339.
- Riederer, P., Sofic, E., Rausch, W.D., Schmidt, B., Reynolds, G.P., Jellinger, K. and Youdim, M.B. (1989) Transition metals, ferritin, glutathione, and ascorbic acid in parkinsonian brains. *J Neurochem*, 52, 515-520.
- Robbins, A.H. and Stout, C.D. (1989) The structure of aconitase. *Proteins*, 5, 289-312.
- Robbins, A.H., Stout, C.D., Piskiewicz, D., Gawron, O., Yoo, C.S., Wang, B.C. and Sax, M. (1982) Single crystals of the iron-sulfur enzyme aconitase. *J Biol Chem*, 257, 9061-9063.
- Robert, M.C., Provost, K. and Lefaucheur, F. (1992) Crystallization in gels and related methods. In *Crystallization of nucleic acids and proteins*, A. Ducruix et R.Giege, eds (Oxford: Oxford Univ. Press).
- Robson, K.J., Lehmann, D.J., Wilmhurst, V.L., Livesey, K.J., Combrinck, M., Merryweather-Clarke, A.T., Warden, D.R. and Smith, A.D. (2004) Synergy between the C2 allele of transferrin and the C282Y allele of the haemochromatosis gene (HFE) as risk factors for developing Alzheimer's disease. *J Med Genet*, 41, 261-265.
- Rogers, J.T., Randall, J.D., Cahill, C.M., Eder, P.S., Huang, X., Gunshin, H., Leiter, L., McPhee, J., Sarang, S.S., Utsuki, T., Greig, N.H., Lahiri, D.K., Tanzi, R.E., Bush, A.I., Giordano, T. and Gullans, S.R. (2002) An iron-responsive element type II in the 5'-untranslated region of the Alzheimer's amyloid precursor protein transcript. *J Biol Chem*, 277, 45518-45528.
- Rolfs, A., Kvietikova, I., Gassmann, M. and Wenger, R.H. (1997) Oxygen-regulated transferrin expression is mediated by hypoxia-inducible factor-1. *J Biol Chem*, 272, 20055-20062.
- Rose, I.A. and O'Connell, E.L. (1967) Mechanism of Aconitase Action. I. THE HYDROGEN TRANSFER REACTION. *J. Biol. Chem.*, 242, 1870-1879.

- Rossmann, M.G. (1972) The molecular replacement method. Gordon et Breach, eds. (New York).
- Rossmann, M.G. and Blow, D.M. (1962) The detection of sub-units within the crystallographic asymmetric unit. *Acta Cryst.*, 15, 24.
- Rotig, A., de Lonlay, P., Chretien, D., Foury, F., Koenig, M., Sidi, D., Munnich, A. and Rustin, P. (1997) Aconitase and mitochondrial iron-sulphur protein deficiency in Friedreich ataxia. *Nat Genet*, 17, 215-217.
- Rottkamp, C.A., Raina, A.K., Zhu, X., Gaier, E., Bush, A.I., Atwood, C.S., Chevion, M., Perry, G. and Smith, M.A. (2001) Redox-active iron mediates amyloid-beta toxicity. *Free Radic Biol Med*, 30, 447-450.
- Rouault, T.A. (2002) Post-transcriptional regulation of human iron metabolism by iron regulatory proteins. *Blood Cells Mol Dis*, 29, 309-314.
- Rouault, T.A., Haile, D.J., Downey, W.E., Philpott, C.C., Tang, C., Samaniego, F., Chin, J., Paul, I., Orloff, D., Harford, J.B. and et al. (1992) An iron-sulfur cluster plays a novel regulatory role in the iron-responsive element binding protein. *Biometals*, 5, 131-140.
- Roussel, A. and Cambillaud, C. (1989) Turbo-Frodo. Silicon Graphics, Mountain View, CA, 77.

S

- Schalinske, K.L., Chen, O.S. and Eisenstein, R.S. (1998) Iron differentially stimulates translation of mitochondrial aconitase and ferritin mRNAs in mammalian cells. Implications for iron regulatory proteins as regulators of mitochondrial citrate utilization. *J Biol Chem*, 273, 3740-3746.
- Schiltz, M., Kvik, Å., Svensson, O.S., Shepard, W., de La Fortelle, E., Prangé, T., Kahn, R., Bricogne, G. and Fourme, R. (1997) Protein Crystallography at Ultra-Short Wavelengths: Feasibility Study of Anomalous-Dispersion Experiments at the Xenon K-edge. *J. Synchrotron Rad.*, 4, 287-297.
- Schipper, H.M., Liberman, A. and Stopa, E.G. (1998) Neural heme oxygenase-1 expression in idiopathic Parkinson's disease. *Exp Neurol*, 150, 60-68.
- Schlegl, J., Gegout, V., Schlager, B., Hentze, M.W., Westhof, E., Ehresmann, C., Ehresmann, B. and Romby, P. (1997) Probing the structure of the regulatory region of human transferrin receptor messenger RNA and its interaction with iron regulatory protein-1. *Rna*, 3, 1159-1172.
- Schloss, J.V., Emptage, M.H. and Cleland, W.W. (1984) pH profiles and isotope effects for aconitases from *Saccharomycopsis lipolytica*, beef heart, and beef liver. α -Methyl-cis-aconitate and threo-Ds- α -methylisocitrate as substrates. *Biochemistry*, 23, 4572-4580.
- Schneider, T.R. and Sheldrick, G.M. (2002) Substructure solution with SHELXD. *Acta Cryst.*, D58, 1772-1779.
- Seiser, C., Teixeira, S. and Kühn, L.C. (1993) Interleukin-2-dependent transcriptional and post-transcriptional regulation of transferrin receptor mRNA. *J Biol Chem*, 268, 13074-13080.

Semenza, G.L. (2000) Expression of hypoxia-inducible factor 1: mechanisms and consequences. *Biochem Pharmacol*, 59, 47-53.

Sheldrick, G.M. (1990) Phase annealing in SHELX-90: direct methods for larger structures. *Acta Cryst.*, A46, 467-473.

Sheldrick, G.M. (1997) Patterson superposition an ab initio phasing. *Methods Enzymol*, 276, 628-641.

Sim, G.A. (1959) The distribution of phase angles for structures containing heavy atoms. II. A modification of the normal heavy-atom method for non-centrosymmetrical structures. *Acta Cryst.*, 12, 813-815.

Smith, A.G., Clothier, B., Robinson, S., Scullion, M.J., Carthew, P., Edwards, R., Luo, J., Lim, C.K. and Toledano, M. (1998) Interaction between iron metabolism and 2,3,7,8-tetrachlorodibenzo-p-dioxin in mice with variants of the Ahr gene: a hepatic oxidative mechanism. *Mol Pharmacol*, 53, 52-61.

Smith, J.L. and Thompson, A. (1998) Reactivity of selenomethionine--dents in the magic bullet? *Structure*, 6, 815-819.

Snowdon, D.A. (2003) Healthy aging and dementia: findings from the Nun Study. *Ann Intern Med*, 139, 450-454.

Staples, C.R., Ameyibor, E., Fu, W., Gardet-Salvi, L., Stritt-Etter, A.L., Schurmann, P., Knaff, D.B. and Johnson, M.K. (1996) The function and properties of the iron-sulfur center in spinach ferredoxin: thioredoxin reductase: a new biological role for iron-sulfur clusters. *Biochemistry*, 35, 11425-11434.

Strub, M.P., Hoh, F., Sanchez, J.F., Strub, J.M., Bock, A., Aumelas, A. and Dumas, C. (2003) Selenomethionine and selenocysteine double labeling strategy for crystallographic phasing. *Structure (Camb)*, 11, 1359-1367.

Sturrock, A., Alexander, J., Lamb, J., Craven, C.M. and Kaplan, J. (1990) Characterization of a transferrin-independent uptake system for iron in HeLa cells. *J Biol Chem*, 265, 3139-3145.

T

Tacchini, L., Bianchi, L., Bernelli-Zazzera, A. and Cairo, G. (1999) Transferrin receptor induction by hypoxia. HIF-1-mediated transcriptional activation and cell-specific post-transcriptional regulation. *J Biol Chem*, 274, 24142-24146.

Tacchini, L., Recalcati, S., Bernelli-Zazzera, A. and Cairo, G. (1997) Induction of ferritin synthesis in ischemic-reperfused rat liver: analysis of the molecular mechanisms. *Gastroenterology*, 113, 946-953.

Taketani, S., Kakimoto, K., Ueta, H., Masaki, R. and Furukawa, T. (2003) Involvement of ABC7 in the biosynthesis of heme in erythroid cells: interaction of ABC7 with ferrochelatase. *Blood*, 101, 3274-3280.

Tang, Y. and Guest, J.R. (1999) Direct evidence for mRNA binding and post-transcriptional regulation by *Escherichia coli* aconitases. *Microbiology*, 145 (Pt 11), 3069-3079.

Teixeira, S. and Kühn, L.C. (1991) Post-transcriptional regulation of the transferrin receptor and 4F2 antigen heavy chain mRNA during growth activation of spleen cells. *Eur J Biochem*, 202, 819-826.

Terwilliger, T.C. (2000) Maximum-likelihood density modification. *Acta Cryst.*, D56, 965-972.

Testa, U., Kühn, L., Petrini, M., Quaranta, M.T., Pelosi, E. and Peschle, C. (1991) Differential regulation of iron regulatory element-binding protein(s) in cell extracts of activated lymphocytes versus monocytes-macrophages. *J Biol Chem*, 266, 13925-13930.

Thayer, M.M., Ahern, H., Xing, D., Cunningham, R.P. and Tainer, J.A. (1995) Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. *Embo J*, 14, 4108-4120.

Theil, E.C. (2000) Targeting mRNA to regulate iron and oxygen metabolism. *Biochem Pharmacol*, 59, 87-93.

Thompson, K., Menzies, S., Muckenthaler, M., Torti, F.M., Wood, T., Torti, S.V., Hentze, M.W., Beard, J. and Connor, J. (2003) Mouse brains deficient in H-ferritin have normal iron concentration but a protein profile of iron deficiency and increased evidence of oxidative stress. *J Neurosci Res*, 71, 46-63.

Tickle, I.J., Laskowski, R.A. and Moss, D.S. (1998) Rfree and the rfree ratio. I. Derivation of expected values of cross-validation residuals used in macromolecular least-squares refinement. *Acta Crystallogr D Biol Crystallogr*, 54, 547-557.

Torti, F.M. and Torti, S.V. (2002) Regulation of ferritin genes and protein. *Blood*, 99, 3505-3516.

Toth, I., Yuan, L., Rogers, J.T., Boyce, H. and Bridges, K.R. (1999) Hypoxia alters iron-regulatory protein-1 binding capacity and modulates cellular iron homeostasis in human hepatoma and erythroleukemia cells. *J Biol Chem*, 274, 4467-4473.

Trenor, C.C., 3rd, Campagna, D.R., Sellers, V.M., Andrews, N.C. and Fleming, M.D. (2000) The molecular defect in hypotransferrinemic mice. *Blood*, 96, 1113-1118.

U

Uversky, V.N., Li, J. and Fink, A.L. (2001) Metal-triggered structural transformations, aggregation, and fibrillation of human alpha-synuclein. A possible molecular link between Parkinson's disease and heavy metal exposure. *J Biol Chem*, 276, 44284-44296.

V

Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.*, 30, 1022-1025.

Van Landeghem, G.F., Sikstrom, C., Beckman, L.E., Adolfsson, R. and Beckman, L. (1998) Transferrin C2, metal binding and Alzheimer's disease. *Neuroreport*, 9, 177-179.

Vernede, X. and Fontecilla-Camps, J.C. (1999) A method to stabilize reduced and/or gas-treated protein crystals by flash-cooling under a controlled atmosphere. *Journal of Applied Crystallography*, 32, 505-509.

Vidal, R., Ghetti, B., Takao, M., Brefel-Courbon, C., Uro-Coste, E., Glazier, B.S., Siani, V., Benson, M.D., Calvas, P., Miravalle, L., Rascol, O. and Delisle, M.B. (2004) Intracellular ferritin accumulation in neural and extraneural tissue characterizes a neurodegenerative disease associated with a mutation in the ferritin light polypeptide gene. *J Neuropathol Exp Neurol*, 63, 363-380.

Vile, G.F. and Tyrrell, R.M. (1993) Oxidative stress resulting from ultraviolet A irradiation of human skin fibroblasts leads to a heme oxygenase-dependent increase in ferritin. *J Biol Chem*, 268, 14678-14681.

Volbeda, A. (1999) Spéleologie des hydrogénase à nickel et à fer. *Les Ecoles Physique et Chimie du Vivant*, 1, 47-52.

Volbeda, A., Charon, M.H., Piras, C., Hatchikian, E.C., Frey, M. and Fontecilla-Camps, J.C. (1995) Crystal structure of the nickel-iron hydrogenase from *Desulfovibrio gigas*. *Nature*, 373, 580-587.

Vulpe, C.D., Kuo, Y.M., Murphy, T.L., Cowley, L., Askwith, C., Libina, N., Gitschier, J. and Anderson, G.J. (1999) Hephaestin, a ceruloplasmin homologue implicated in intestinal iron transport, is defective in the *sla* mouse. *Nat Genet*, 21, 195-199.

W

Wang, B.C. (1985) Resolution of phase ambiguity in macromolecular crystallography. *Methods in enzymology*, 115, 90-111.

Wang, J., Chen, G., Muckenthaler, M., Galy, B., Hentze, M.W. and Pantopoulos, K. (2004) Iron-mediated degradation of IRP2, an unexpected pathway involving a 2-oxoglutarate-dependent oxygenase activity. *Mol Cell Biol*, 24, 954-965.

Wang, J. and Pantopoulos, K. (2002) Conditional derepression of ferritin synthesis in cells expressing a constitutive IRP1 mutant. *Mol Cell Biol*, 22, 4638-4651.

Wardrop, S.L. and Richardson, D.R. (1999) The effect of intracellular iron concentration and nitrogen monoxide on Nramp2 expression and non-transferrin-bound iron uptake. *Eur J Biochem*, 263, 41-49.

Wardrop, S.L., Watts, R.N. and Richardson, D.R. (2000) Nitrogen monoxide activates iron regulatory protein 1 RNA-binding activity by two possible mechanisms: effect on the [4Fe-4S] cluster and iron mobilization from cells. *Biochemistry*, 39, 2748-2758.

Waters, M.L. (2002) Aromatic interactions in model systems. *Curr Opin Chem Biol*, 6, 736-741.

Weinstein, D.A., Roy, C.N., Fleming, M.D., Loda, M.F., Wolfsdorf, J.I. and Andrews, N.C. (2002) Inappropriate expression of hepcidin is associated with iron refractory anemia: implications for the anemia of chronic disease. *Blood*, 100, 3776-3781.

Weiss, G., Goossen, B., Doppler, W., Fuchs, D., Pantopoulos, K., Werner-Felmayer, G., Wachter, H. and Hentze, M.W. (1993) Translational regulation via iron-responsive elements by the nitric oxide/NO-synthase pathway. *Embo J*, 12, 3651-3657.

Weiss, G., Werner-Felmayer, G., Werner, E.R., Grunewald, K., Wachter, H. and Hentze, M.W. (1994) Iron regulates nitric oxide synthase activity by controlling nuclear transcription. *J Exp Med*, 180, 969-976.

Williams, C.H., Stillman, T.J., Barynin, V.V., Sedelnikova, S.E., Tang, Y., Green, J., Guest, J.R. and Artymiuk, P.J. (2002) E. coli aconitase B structure reveals a HEAT-like domain with implications for protein-protein recognition. *Nat Struct Biol*, 9, 447-452.

Wilms, H., Rosenstiel, P., Sievers, J., Deuschl, G., Zecca, L. and Lucius, R. (2003) Activation of microglia by human neuromelanin is NF-kappaB dependent and involves p38 mitogen-activated protein kinase: implications for Parkinson's disease. *Faseb J*, 17, 500-502.

Wlodawer, A. and Hodgson, K.O. (1975) Crystallization and drystal data for monellin. *Proc. Nat. Acad. Sc. USA*, 72, 398-399.

Wu, K.J., Polack, A. and Dalla-Favera, R. (1999) Coordinated regulation of iron-controlling genes, H-ferritin and IRP2, by c-MYC. *Science*, 283, 676-679.

Y

Yamanaka, K., Ishikawa, H., Megumi, Y., Tokunaga, F., Kanie, M., Rouault, T.A., Morishima, I., Minato, N., Ishimori, K. and Iwai, K. (2003) Identification of the ubiquitin-protein ligase that recognizes oxidized IRP2. *Nat Cell Biol*, 5, 336-340.

Yang, J., Goetz, D., Li, J.Y., Wang, W., Mori, K., Setlik, D., Du, T., Erdjument-Bromage, H., Tempst, P., Strong, R. and Barasch, J. (2002) An iron delivery pathway mediated by a lipocalin. *Mol Cell*, 10, 1045-1056.

Yeates, T.O. (1997) Detecting and Overcoming Crystal Twinning. In Carter, C.W.J. and Sweet, R.M. (eds.), *Macromolecular Crystallography. Part A*. Academic Press, San Diego et al., Vol. 276, pp. 344-358.

Yoshida, K., Furihata, K., Takeda, S., Nakamura, A., Yamamoto, K., Morita, H., Hiyamuta, S., Ikeda, S., Shimizu, N. and Yanagisawa, N. (1995) A mutation in the ceruloplasmin gene is associated with systemic hemosiderosis in humans. *Nat Genet*, 9, 267-272.

Z

Zambenedetti, P., De Bellis, G., Biunno, I., Musicco, M. and Zatta, P. (2003) Transferrin C2 variant does confer a risk for Alzheimer's disease in caucasians. *J Alzheimers Dis*, 5, 423-427.

- Zecca, L., Gallorini, M., Schunemann, V., Trautwein, A.X., Gerlach, M., Riederer, P., Vezzoni, P. and Tampellini, D. (2001) Iron, neuromelanin and ferritin content in the substantia nigra of normal subjects at different ages: consequences for iron storage and neurodegenerative processes. *J Neurochem*, 76, 1766-1773.
- Zecca, L., Youdim, M.B., Riederer, P., Connor, J.R. and Crichton, R.R. (2004) Iron, brain ageing and neurodegenerative disorders. *Nat Rev Neurosci*, 5, 863-873.
- Zheng, L., Kennedy, M.C., Beinert, H. and Zalkin, H. (1992) Mutational analysis of active site residues in pig heart aconitase. *J Biol Chem*, 267, 7895-7903.
- Zhou, B., Westaway, S.K., Levinson, B., Johnson, M.A., Gitschier, J. and Hayflick, S.J. (2001) A novel pantothenate kinase gene (PANK2) is defective in Hallervorden-Spatz syndrome. *Nat Genet*, 28, 345-349.
- Zoghbi, H.Y. and Orr, H.T. (2000) Glutamine repeats and neurodegeneration. *Annu Rev Neurosci*, 23, 217-247.